AD-A239 946

(2)

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DTIC
ELECTE
AUG 27 1991
S B D

THESIS

ESTIMATING COMMUTE DISTANCES OF U.S. ARMY RESERVISTS BY REGIONAL AND UNIT CHARACTERISTICS

by

Steven E. Galing

September 1990

Thesis Advisor: Laura Johnson

Approved for public release; distribution is unlimited

91-08869

91 8 26 011

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b OFFICE SYMBOL (If applicable) Code 30	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	7b ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)
ESTIMATING COMMUTE DISTANCES OF U.S. ARMY RESERVISTS BY REGIONAL AND UNIT CHARACTERISTICS

12. PERSONAL AUTHOR(S)
Galing, Steven E.

13a. TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Master's Thesis	FROM ____ TO ____	1990 September	120

16 SUPPLEMENTARY NOTATION The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Euclidean distance, Power Transformation, Maximum Likelihood Estimator, Multiple Regression

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This thesis develops a multiple regression model using regional and unit characteristics to estimate commuting distances of U.S. Army Reservists. The data were obtained from a 1988 file established by the Defense Manpower Data Center containing locational and biodemographic information on 238,174 enlisted reservists. A random sample of 91 reserve centers was selected for the analysis.

The logistic and normal distributions were evaluated as possible candidates for fitting the commuting distance distribution. It was found that a power transformation of the fractional distance traveled fit both distributions quite well. Parameters for the two distributions are obtained through a method of maximum likelihood estimation. Finally,

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Laura Johnson	22b TELEPHONE (Include Area Code) (408) 646-2569 22c OFFICE SYMBOL OR/Jo

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603 Unclassified

#19 Continued:

a multiple regression equation is used to estimate the parameters of the commute distance distribution as a function of reserve center and market characteristics.

The results of the multiple regression equation provide the U.S. Army Recruiting Command with some important variables necessary to predict commuting distances.

Approved for public release; distribution is unlimited

Estimating Commute Distances of U.S. Army Reservists by Regional and Unit Characteristics

by

Steven E. Galing
Captain, United States Army
B.S., United States Military Academy, 1980
M.B.A., Western New England College, 1988

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1990

Author: Steven E. Galing

Approved by: Laura Johnson, Thesis Advisor

George Thomas, Second Reader

Peter Purdue, Chairman,
Department of Operations Research

ABSTRACT

This thesis develops a multiple regression model using regional and unit characteristics to estimate commuting distances of U.S. Army Reservists. The data were obtained from a 1988 file established by the Defense Manpower Data Center containing locational and biodemographic information on 238,174 enlisted reservists. A random sample of 91 reserve centers was selected for the analysis.

The logistic and normal distributions were evaluated as possible candidates for fitting the commuting distance distribution. It was found that a power transformation of the fractional distance traveled fit both distributions quite well. Parameters for the two distributions are obtained through a method of maximum likelihood estimation. Finally, a multiple regression equation is used to estimate the parameters of the commute distance distribution as a function of reserve center and market characteristics.

The results of the multiple regression equation provide the U.S. Army Recruiting Command with some important variables necessary to predict commuting distances.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. GENERAL

The Department of Defense's share of the national budget may soon be decreasing. As communism in Europe and Asia continues to crumble, funds may be reallocated from the Armed Forces to offset the national debt and bolster domestic and foreign expenditures. The U.S. Army is preparing for possible funding and personnel reductions.

The Secretary of Defense has proposed four major reductions in the 1991 Army budget that will save an estimated $3.3 billion. These proposals are shown in Table 1 [Ref. 1:p. 19].

TABLE 1

PROPOSED REDUCTIONS IN THE ARMY'S 1991 BUDGET

Proposal	Cost Savings (in billions)
1. Deactivate 2 Army Divisions	$1.2
2. Cut Helicopter Improvement Program	0.3
3. End Production of the M-1 Tank	1.1
4. Cancel Procurement of the Apache Helicopter	0.7
	Total: $3.3

These sweeping proposals will shift a higher percentage of the Army's total defensive capability to the reserves. Currently, there are 18 active and 12 reserve divisions; 40 percent of the Army's fighting force is in the reserves. If the Secretary of Defense's budget proposal is approved, the force mix will be 16 active and 12 reserve by the end of 1991. The reserves will then comprise over 42 percent of the Army's defensive capability. Future possible reductions in 1992 and beyond foretell of increased reliance on the reserves. In short, the Army Reserve will continue to be an integral part of this country's defense forces.

In light of the impending reductions, the reserves may expect to be affected in one of four ways. First, only active forces might be reduced, greatly increasing the current 40 percent ratio of reserve to total Army units. Secondly, deactivated units might be transferred to the reserves, swelling the reserve strength and necessitating the opening of new reserve centers. Thirdly, the reserves will undergo troop reductions independent of active cutbacks. Lastly, the ratio of reserve units in the total Army force could remain at approximately 40 percent, implying that the reserves will suffer proportionally the same reductions as the active Army. While it is difficult to predict which option, if any, will become reality, one can argue that the Army must be prepared either to establish new reserve centers or close down existing ones.

Manning the reserve units is extremely important. A unit cannot be expected to perform its mission unless it has trained, qualified personnel staffing critical positions. Units are considered undeployable if they cannot attain certain fill rates.

The U.S. Army Recruiting Command (USAREC) is the proponent agency responsible for recruiting personnel to meet unit fill rates. An important objective of the USAREC is to identify market areas that will support new reserve units [Ref. 2:p. 5]. Most unit location studies have centered on supply variables such as civilian wages and unemployment [Ref. 2:p. 11]. Little or no research has been done to establish a link between various supply variables and commuting distances.

B. PURPOSE

The purpose of this thesis is to analyze individual commuting distances of enlisted reservists and estimate a statistical relationship between commute behavior and different supply variables. Such an estimated relationship can be used by the USAREC to analyze a geographical area and assess how it might support a new reserve unit.

This thesis utilizes the logistic and normal distributions with suitable power transformations to model commuting patterns. The fitted distributions can then be used by the USAREC to determine the geographic extent of a recruiting market around a particular reserve center.

C. BACKGROUND

Commuting patterns have changed significantly over the past 100 years. In the period 1890 - 1920, people primarily commuted to the Central Business District (CBD). The main source of transport was the trolley [Ref. 3:p. 512]. The average commute distance using the trolley was 1.8 miles [Ref. 4:p. 361]. For convenience, houses were located within a short walking distance of the CBD or the trolley stop [Ref. 3:p. 512].

After 1920, the introduction of the bus and private automobile increased commuting distances significantly (See Figure 1).

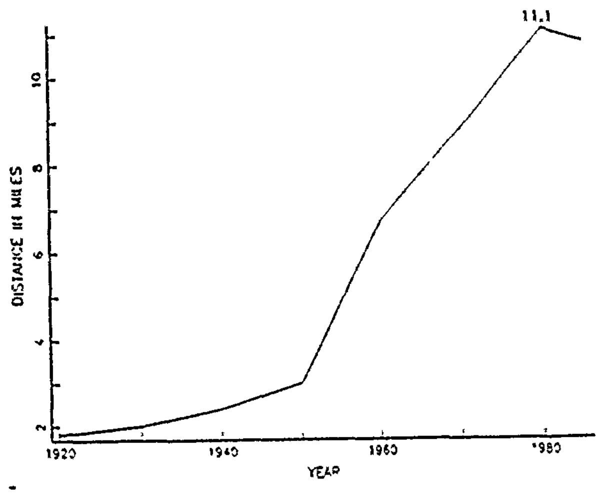

Figure 1

Commute Distances: 1920-1985

People had the freedom to disperse from transit lines, allowing cities to expand spatially. In the 1960's and 1970's an extensive highway and interstate system was developed to accommodate the dispersed population [Ref 4:p. 361]. Businesses began to move manufacturing and service sites out of the CBD to more suburban locations [Ref. 3:p. 515]. Because many businesses moved to the suburbs, commuting distances peaked at 11.1 miles in 1979 [Ref. 5:p. 3]. The trend for businesses to move from the CBD to the suburbs continued into the 1980's, as the average commuting distance decreased to 10.8 miles in 1985 [Ref. 6]. Therefore, one can infer that in the future commute distances of employees will stay the same or decrease slightly.

Distances that personnel are willing to commute should be considered when assessing potential and existing Army Reserve Center locations. Research using commuting distances to predict manpower supply for the Naval Reserves was done by Beth Asch in 1985 [Ref. 7]. She contends that potential reservists will commute at most 100 miles. Therefore, she recommends that the Navy assess the geographical region within 100 miles from a potential site to determine if the area can provide the necessary personnel to staff the ship(s). Likewise, the Army can use commuting distances to assess the feasibility of a geographical region to staff reserve units.

A file containing location and biodemographic data on 238,174 enlisted reservists established in 1988 by the Defense

Manpower Data Center was used to develop this thesis. The individual commuting distances were estimated by using longitudinal and latitudinal data to obtain the Euclidean distance from the individual's home zip code to the unit's zip code. Therefore, the distances used in this thesis are estimates and not the actual commuting distances. Hence, one must assume that the relationship between the estimated distance and the true distance remains the same within a reserve center all else being equal.

D. ASSUMPTIONS

Listed below are assumptions that were made prior to analyzing the data.

1. Reserve centers are located at the centroid of the zip code region.

2. Reservists reside uniformly throughout a zip code region. This is important because distances were computed by measuring the Euclidean distance from the centroid of the reserve center zip code region to the centroid of the reservists' zip code.

3. The average person will commute at most 100 miles. Therefore, commute distances in excess of 100 miles are not considered in this analysis.

4. Commute distances of personnel residing in the same zip code region as the reserve center are greater than zero miles. This is important because the Euclidean distances in the database for these personnel equal zero miles. Therefore, through the use of a detailed road map, zip code boundaries were estimated and the distance across the zip code area was multiplied by 0.5 (0.5 was selected because this is the farthest possible commute distance to the reserve center). For example, if the distance across a zip code is 6 miles, then personnel residing in that zip code can travel at most 3.0 miles.

5. Commute distances will vary from one region of the country to another. Therefore, 91 reserve centers were

selected using a random number generator from the Northeast, South, Midwest, and West. The number sampled from each region is proportional to the personnel in the reserves from that region. For example, since approximately 27 percent of reservists are located in the West, 25 of the 91 reserve centers (27 percent) are from the West.

E. EXPLANATION OF DATA AND RESEARCH

Appendix A contains Tables 2-5 listing the reserve centers selected in each geographical region. Some personnel did not have a home zip code, so they are not considered in the analysis. Also shown in each table are the number of personnel that commuted over 100 miles and the number living in the reserve center zipcode. The necessary distance adjustments discussed in assumption number four are shown in Tables 6-9 in Appendix B. One should note that no more than 10 percent of the reservists in a geographical region reside in the same zip code area as the reserve center. In addition, the analysis includes at least 91 percent of the distance data within 100 miles of each reserve center.

The next step is to evaluate the logistic and normal distributions and determine whether commuting distances can be accurately predicted. Because the data are not symmetric, a suitable power transformation is applied. Parameters for the logistic and normal distributions, in addition to the power transformation, are obtained through a method of maximum likelihood estimation using a numerical solver. The estimated parameters are substituted into the cumulative distribution functions to compute the appropriate percentile commute distance. Finally, regression analysis techniques are investigated to determine if regional and unit characteristics can be used to predict the distribution parameters. These steps will be discussed in detail in the next three chapters.

II. LOGISTIC DISTRIBUTION

A. GENERAL

Research characterizing part-time commute behavior for Army reservists was done by Laura D. Johnson and George W. Thomas using 30 reserve centers [Ref. 8]. Sampling from the 1988 reserve data base, they conducted extensive exploratory data analysis to determine if there exists a probability distribution that adequately describes some power transformation of the commuting distances. Relying primarily on the Kolmogorov-Smirnov statistic, they determined that the logistic distribution provided a good fit [Ref. 8:p. 7]. Based on their results, the power logistic distribution was initially selected as a model of commuting distances of U.S. Army Reservists.

The logistic distribution has previously been used as a growth curve and for demographic purposes [Ref. 9:p. 3]. It is symmetric and has a shape similar to the normal distribution. The logistic has relatively "longer tails" meaning that it has greater variability about the mean [Ref. 9:p. 6]. For example, the inflection points on the standard normal curve are at ± 1 while the logistic distribution has its inflection points at ± 0.53 [Ref. 9:p. 6]. The relatively "longer tails" have their most significant impact on the fourth central moment. The value of the logistic is 4.2 vice 3.0 for the normal distribution [Ref. 9:p. 6]. Although the relatively "longer tails" have a considerable effect on the fourth central moment, there is a much smaller effect on the estimation of specific values of the cumulative distribution function (CDF). Thus, one could argue that it is appropriate to use either the logistic or normal distribution to predict the 75th and 90th percentile commute distances. Shown below is the logistic distribution [Ref. 9].

$$f(y) = \frac{e^{-\frac{(y_i-\alpha)}{\beta}}}{\beta\left[1 + e^{-\frac{(y_i-\alpha)}{\beta}}\right]^2}$$

$$\text{where} \quad \begin{aligned} E[y] &= \alpha \\ Var[y] &= \frac{\beta^2\pi^2}{3} \end{aligned}$$

B. POWER TRANSFORMATION

The commuting distances were examined to determine if a power transformation was necessary. One fast and efficient method is to plot the data on a quantile plot. As an example, the commuting distances for the reserve center at Long Beach, California are plotted in Figure 2.

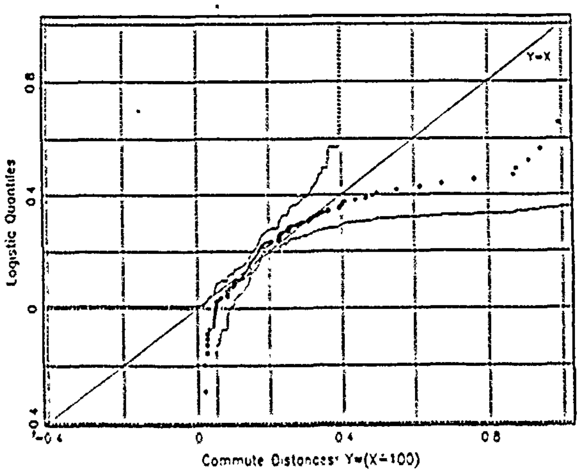

Figure 2

Long Beach, CA   N=151

Logistic Quantile Plot Without a Transformation

If the data do not fall close to the Y=X line, the data are not symmetric and will fail the Kolmogorov-Smirnov statistic. [Ref. 10:p. 30] One way to get the data to lie close to the Y=X line is to apply a suitable power transformation to the commuting distances.

The distance travelled by the ith member of a reserve unit is denoted by x_i . The transformation used is [Ref. 8:p. 7]:

$$y_i = \begin{cases} x_i^{\theta} & if\ \theta \neq 0 \\ \ln x_i & if\ \theta = 0 \end{cases}$$

If the random variable Y = u(X) defines a one-to-one correspondence between the values of X and Y, the equation y = u(x) will yield a unique value for x in terms of y. This relationship can be written as x = w(y). Therefore, the probability distribution of Y is:

$$g(y) = f[w(y)]\ |J|$$

where |J| is the jacobian of the transformation. [Ref. 11:p. 143]

To obtain the Jacobian, one takes the derivative of y with respect to x.

$$\frac{dy}{dx} = \begin{cases} \theta x_i^{\theta-1} & if\ \theta \neq 0 \\ \dfrac{1}{x_i} & if\ \theta = 0 \end{cases}$$

Substituting into the equation for g(y), one obtains the probability density function (pdf) shown below.

$$g(x_i) = \begin{cases} f(x_i^{\theta})\ \theta\ x_i^{\theta-1} & if\ \theta \neq 0 \\ f(\ln x_i)\ \dfrac{1}{x_i} & if\ \theta = 0 \end{cases}$$

These equations will yield the commute distance for each individual after the power transformation.

C. MAXIMUM LIKELIHOOD ESTIMATION

A method of maximum likelihood was used to estimate the parameters of the logistic distribution. A numerical solver called the General Algebraic Modelling System (GAMS) was utilized to obtain values for α, β, and θ (call these values a, b, and t) [Ref. 12]. Each reserve center has n observations. The individual commute distances, x_1, x_2, ..., x_n, are used to obtain the following likelihood for each center [Ref. 13:p. 268].

$$L_\theta(\alpha,\beta) = \prod_{i=1}^{n} g(x_i)$$

The purpose is to find estimates of a, b, and t that have the following desirable property [Ref. 13:p. 268].

$$L_t\ (a,b) \geq L_\theta(\alpha,\beta) \qquad \forall\ estimates\ of\ (\alpha,\beta,\theta)$$

These are called the maximum likelihood estimators (MLE) and are obtained by taking the derivative with respect to a, b, and t and setting each equation equal to zero. Solutions to these equations were obtained numerically in GAMS. The derivations to obtain a, b, and t are shown in Appendix C. Obtaining optimal estimates can be time consuming because the bounds for a, b, and t must be established by trial and error. This process was done to each of the 91 reserve centers and is tabulated in Tables 10-13 in Appendix D. An example of one GAMS program and the output is in Appendix E.

One way to easily verify the accuracy of the GAMS output is to plot the commute distances after applying the calculated power transformation (t). A quantile plot using the reserve center data from Long Beach, California is depicted in Figure 3.

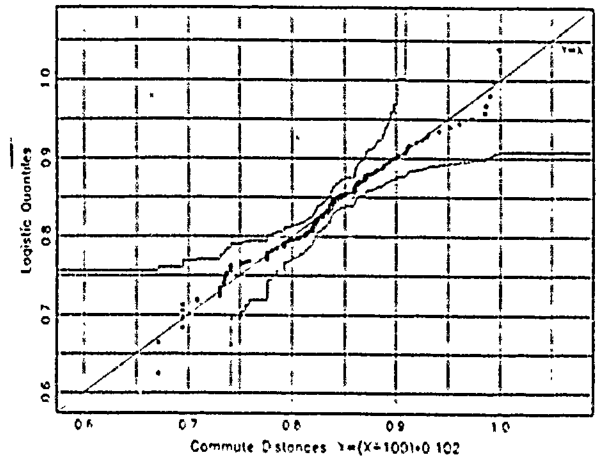

Figure 3

Long Beach, CA N=151

Logistic Quantile Plot With a Transformation

The data lie very close to the Y=X line indicating that the fit is quite accurate. In addition, the transformed data easily pass Komogorov-Smirnov statistic, effectively validating the model.

The a, b, and t values can be substituted into the logistic CDF to obtain estimated commute distances for the 50th, 75th, and 90th percentile commute distances (See

Appendix F for calculations). The estimated commute distances from the logistic distribution versus the actual commute distances are in Tables 14-17 in Appendix G.

D. RESULTS

The power logistic distribution appears to be an accurate predictor of commuting distances. Estimated commute distances for the 50th percentile were very close to the actual distances. Let the prediction error be the actual distance minus the predicted distance. The absolute mean error is 1.8 miles or less for each geographical region. Figure 4 is a box plot of the 50th percentile prediction errors. It is interesting to note that the greatest variability is in the West. Perhaps rural living accounts for this anomaly.

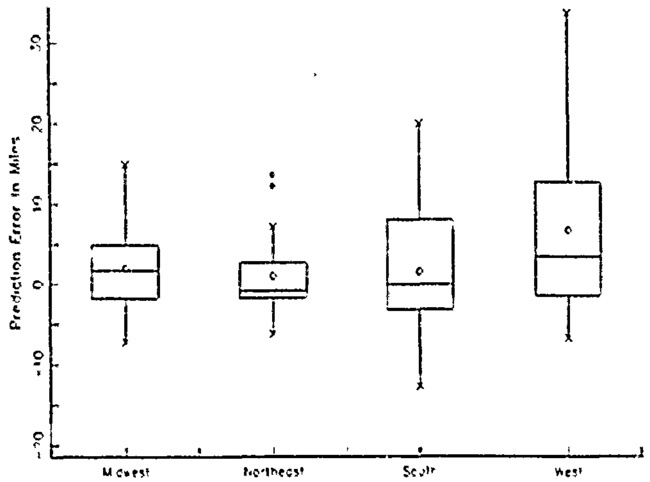

Figure 4

50th Percentile Commute Distances

Errors In The Logistic Predictions

The predicted commuting distances for the 75th percentile are shorter than the actual distances. This is not unexpected because of the "longer tails" discussed earlier. The prediction errors for the 75th percentile are depicted in Figure 5.

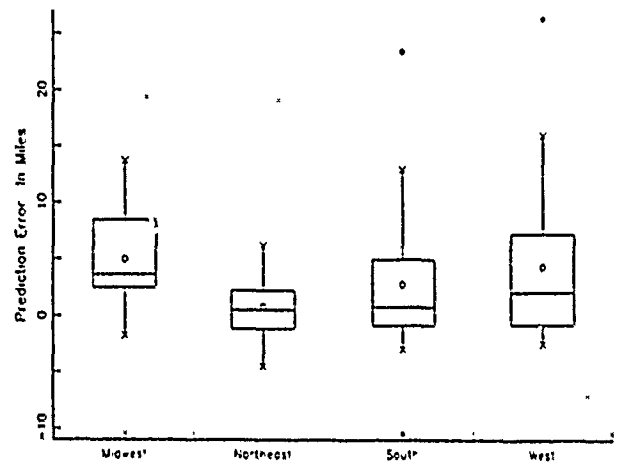

Figure 5

75th Percentile Commute Distances

Errors In The Logistic Predictions

As expected, predictions in the Northeast are more accurate than the other regions. This is probability due to the highly developed, closely located cities in the Northeast. The spread of workers in the South and West and the resulting longer commute distances contribute to the greater variability in these regions.

Predicted commuting distances for the 90th percentile were very accurate in the Midwest and Northeast. Figure 6 indicates that the West has much greater variability than the other three regions.

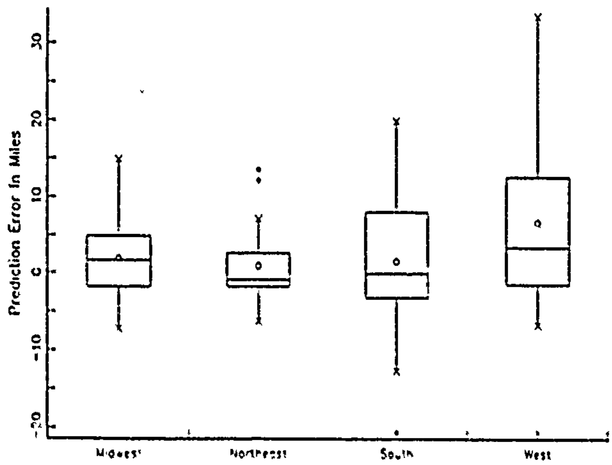

Figure 6

90th Percentile Commute Distances

Errors In The Logistic Predictions

This is understandable because many communities in states such as Nevada and Colorado are long distances from the nearest reserve center. Several commuters traveling a long distance can skew predictions in the model.

Overall, the logistic distribution is an accurate predictor of commuting distance percentiles. Rural living and the dispersion of communities might account for some of the greater variability in the South and West. The more accurate Northeastern predictions may be a result of the highly developed urban environment from Boston, Massachusetts to Washington, D.C. Another possible explanation for the smaller error in the Northeast is that most reserve centers have been purposely located to ensure that the 90th percentile commute distance is less than 50 miles. Predictions, particularly in the South and West, might have been effected by the elimination of distances in excess of 100 miles. Lastly, it is conceivable that only 10% of reservists are willing to commute over 50 miles. One can surmise that if a center is not located within 50 miles of their residence, they seek alternate employment.

III. NORMAL DISTRIBUTION

A. GENERAL

As discussed in Chapter II, the logistic distribution is similar in shape to the normal distribution. Listed below are several characteristics common to both distributions [Ref. 11:p. 112].

1. It is continuous and symmetrical about the mean (μ).
2. The mode equals the mean.
3. The curve has points of inflection at $\mu \pm \sigma$ (σ is the standard deviation).
4. The curve approaches the horizontal axis asymptotically to the left and right of μ.
5. The area under the curve is equal to 1.

Because of their similarities, one can speculate that commute distances under some suitable power transformation could also be modelled using the normal distribution. The hypothesis being tested is that the normal distribution will yield results that are as accurate as the logistic distribution.

The normal distribution is depicted below [Ref. 13:pp. 90-92].

$$f(v_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(v_i-\mu)^2}{2\sigma^2}}$$

$$\text{where} \quad \begin{aligned} E[v_i] &= \mu \\ Var[v_i] &= \sigma^2 \end{aligned}$$

B. POWER TRANSFORMATION

The data were examined prior to applying a transformation. Again, a quantile plot was chosen because it adequately displays all of the data (See Figure 7).

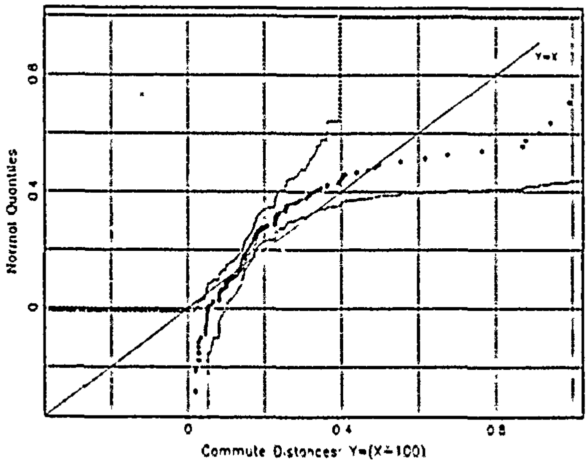

Figure 7

Long Beach, CA N=151

Normal Quantile Plot Without a Transformation

Because the data do not fit the Y=X line, an appropriate power transformation was applied. The transformation selected is shown below [Ref. 14:p. 214].

$$v_i = \begin{cases} \dfrac{x_i^{\theta}-1}{\theta} & if\ \theta \neq 0 \\ \ln x_i & if\ \theta = 0 \end{cases}$$

where x_i is the distance travelled by the ith member of the reserve unit.

This transformation is appropriate because it is continuous at 0 [Ref. 14:p. 215].

As discussed in Chapter II, if the random variable V = u(X) defines a one-to-one correspondence between the values of V and X, the equation v = u(x) will yield a unique value for x in terms of v. This relationship can be written as x = w(v). Therefore, the probability distribution of V is:

$$g(v) = f[w(v)]\ |J|$$

where $|J|$ is the jacobian of the transformation. [Ref. 11:p. 143]

To obtain the Jacobian, one takes the derivative of v with respect to x_i .

$$\frac{dv_i}{dx_i} = \begin{cases} x_i^{\theta-1} & if\ \theta \neq 0 \\ \dfrac{1}{x_i} & if\ \theta = 0 \end{cases}$$

Substituting into the equation for g(v), one obtains the following pdf:

$$g(x_i) = \begin{cases} f\left(\dfrac{x_i^{\theta}-1}{\theta}\right) x_i^{\theta-1} & if\ \theta \neq 0 \\ f\ (\ln x_i)\ \dfrac{1}{x_i} & if\ \theta = 0 \end{cases}$$

Following the same procedures as discussed in the logistic distribution, the likelihood function is obtained:

$$L_\theta(\alpha,\beta) = \prod_{i=1}^{n} g(x_i)$$

$$\alpha = \mu$$

$$\textit{where} \quad \beta = \sigma$$

$$\theta = \textit{power transformation}$$

Substituting for $g(x_i)$ when $\theta \neq 0$, one obtains:

$$L_\theta(\alpha,\beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\left(\frac{x_i^\theta-1}{\theta}-\mu\right)^2}{2\sigma^2}} \quad (x_i^{\theta-1})$$

After taking the natural logarithm of both sides one can simplify the equation to obtain:

$$\ln L_\theta(\alpha,\beta) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\frac{x_i^\theta-1}{\theta}-\mu\right)^2 + \sum_{i=1}^{n}(\theta-1)\ln(x_i)$$

Substituting for $g(x_i)$ when $\theta=0$, one obtains:

$$L_\theta(\alpha,\beta) = \prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(\ln x_i-\mu)^2}{2\sigma^2}}\right)\left(\frac{1}{x_i}\right)$$

One can take the natural logarithm of both sides and obtain:

$$\ln L_\theta(\alpha,\beta) = -\frac{n}{2}\ln 2\pi - n\ln\theta - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2 - \sum_{i=1}^{n}\ln x_i$$

After obtaining the estimated parameters from GAMS, these equations will yield the commute distance distribution for each reserve center.

GAMS was used to implement a method of maximum likelihood to estimate the parameters of the normal distribution. The estimators for α, β, and θ are labelled m, s, and t. An example of one GAMS program and the associated output is in Appendix H. Values for each of the 91 reserve centers were computed in a similar manner and are tabulated in Tables 18-21 in Appendix I.

The GAMS model to obtain the estimates can be verified by graphing the data. After applying the power transformation t, the data were graphed on a quantile plot (Figure 8).

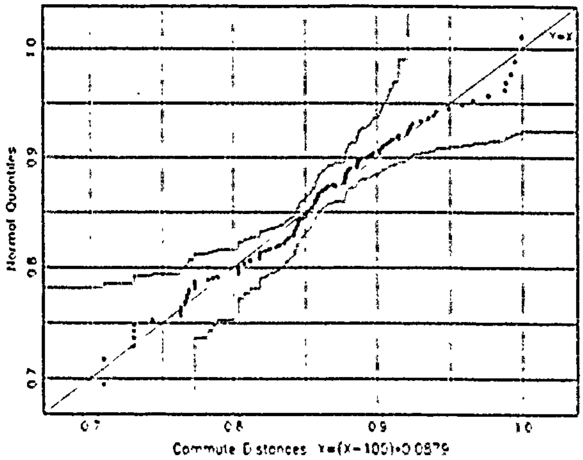

Figure 8

Long Beach, CA N=151

Normal Quantile Plot With a Transformation

Clearly, the transformed data lie very close to the Y=X line. In addition, the Kolmogorov-Smirnov statistic is satisfied providing further proof that the transformation is accurate.

The values of m, s, and t are substituted into the normal CDF to obtain estimated commute distances for the 50th, 75th, and 90th percentiles. Calculations are in Appendix J. The estimated commute distances are in Tables 22-25 in Appendix K.

C. RESULTS

The normal distribution with the appropriate power transformation is an excellent predictor of commuting distances. Figure 9 is a box plot of the 50th percentile prediction errors.

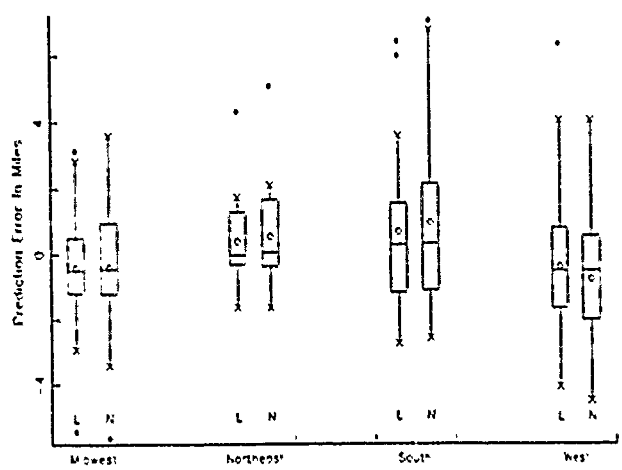

Figure 9

50th Percentile Commute Distances

Errors In The Logistic (L) And Normal (N)

The largest absolute mean error is 2.1 miles in the South. Prediction variability is greatest in the South, a phenomenon that was also observed with the logistic distribution. Estimators are skewed right in the Northeast and South but are skewed left in the West. One possible explanation for predicted distances exceeding actual distances in the west is the geographical dispersion of communities. Many western communities are close together because of the terrain and water availability. Therefore, several miles may separate clusters of communities. Consequently, the commuting flow consists of many different groups of personnel travelling almost the same distance.

Predicted commuting distances for the 75th percentile are depicted in Figure 10.

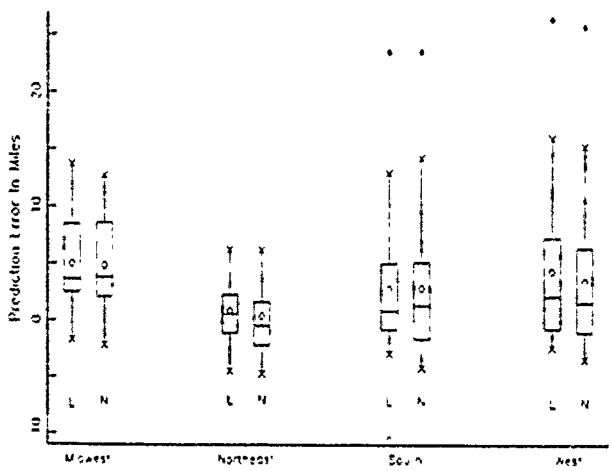

Figure 10

75th Percentile Commute Distances

Errors In The Logistic (L) And Normal (N)

As expected, prediction errors in the Northeast are less than other regions. Although variability is greater in the West, the mean absolute error is only 4.7 miles. This relatively small error indicates that the normal distribution provides a good estimate. The mean absolute error is 5.2 miles in the Midwest. Even though this is higher than the error in the West, the variability is much less, implying that the normal distribution will provide relatively consistent predictions.

Variability and skewness in the 90th percentile predictions for each region are similar to predictions for the 75th percentile. Figure 11 shows that predictions are quite accurate in the Northeast and Midwest.

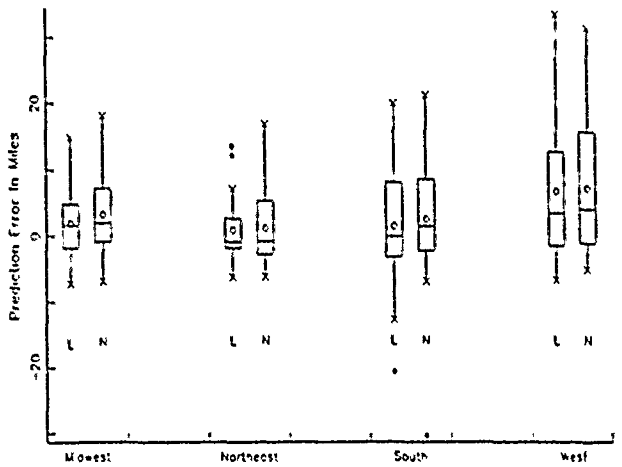

Figure 11

90th Percentile Commute Distances

Errors In The Logistic (L) And Normal (N)

Rural living and the associated spatial movement away from large metropolitan areas contributes to the greater variability in the South and West.

D. COMPARISON OF THE NORMAL AND LOGISTIC DISTRIBUTIONS

The normal and logistic distributions with suitable transformations have been used to predict commute distances. Figure 12 depicts prediction errors for the 50th, 75th, and 90th percentile commute distances for 91 reserve centers.

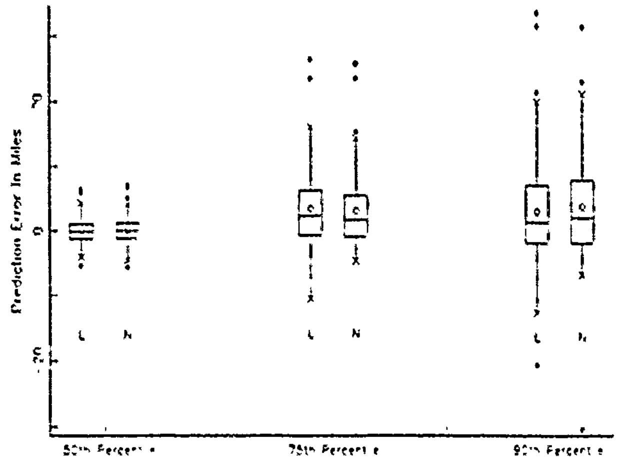

Figure 12

Percentile Commute Distances

Errors In The Logistic (L) And Normal (N)

The 50th percentile predictions are nearly identical for both distributions. A slight difference in variability between the distributions is evident in the 75th percentile. Both distributions are skewed right, with the normal yielding more consistent predictions. The 90th percentile predictions for both distributions are nearly identical. It appears that the "longer tails" of the logistic distribution have little effect when compared to the normal distribution. Because the predictions from both distributions are nearly identical, one can select either distribution to estimate commuting distances.

Although either distribution can adequately forecast commuting distances, the normal distribution is preferred for several reasons. The MLE's for the normal distribution are faster to compute in GAMS than the MLE's for the logistic distribution. Secondly, the MLE's for the normal distribution are asymptotically normally distributed. [Ref. 13:p. 272]. In order to verify that the MLE's possess normal distribution properties, quantile plots with Kolmogorov-Smirnov bounds were evaluated for m, s, and t. Shown below in Figure 13 is the quantile plot of the t values.

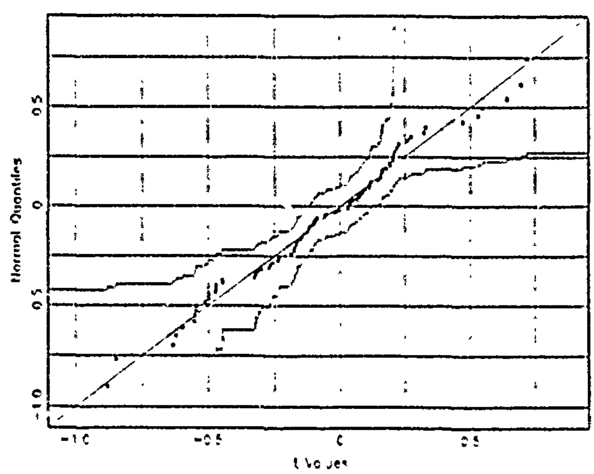

Figure 13

Quantile Plot of t Values

The values clearly lie along the Y = X line and are well inside the Kolmogorov-Smirnov bounds, indicating that the t values are normally distributed. Similar quantile plots were evaluated with identical results for the m and s values. Thus, the MLE's are consistent with the normal distribution. Conversely, the MLE's for the logistic distribution may not be asymptotically efficient because the transformation is discontinuous when $\theta = 0$. In summary, because the computed power normal distribution is continuous and faster to compute it is preferred over the power logistic distribution.

IV. MULTIPLE REGRESSION

A. GENERAL

The objective of this chapter is to establish a multiple regression model using local geographical and unit characteristics that will best predict the values of μ, σ, and θ, based on the estimators m, s, and t found in Chapter III. Listed below are geographical and unit characteristics that were tested as explanatory variables in the regression models.

1. Geographical Characteristics

a. The number of males in category III-A[1] and above from 17-29 years old residing in the reserve center county.

b. The median household income in the reserve center county divided by the average annual military wage at the reserve center.

c. Unemployment rate in the reserve center county.

d. The percentage of full time civilian workers that commute to work outside of their residential county. This variable is a surrogate measurement for the propensity to commute.

2. Unit Level Characteristics

a. The percentage of male personnel stationed at the reserve center.

b. The number of prior service personnel (both men and women) serving at the reserve center.

---

[1]Potential recruits that are classified in category III-A or higher must meet the following educational and testing requirements: High School Degree (HSDG) or Currently in High School past the 10th grade (CIHS) or Alternative Credential Holder (ACH) and scores of at least 50 on qualifying Armed Services Vocational Aptitude Battery Tests (ASVAB) [Ref. 15 :p.14].

In addition to the geographical characteristics, the model was tested by partitioning the United States into nine census regions. The census regions and the states assigned to each region are listed in Appendix L. The model was also tested by unit type. Each reserve center was categorized as combat, combat support, combat service support, or medical based on the largest type of unit assigned to the reserve center. A listing of the different types of units and which of the four categories they were assigned to is in Appendix M.

B. ASSUMPTIONS

Prior to developing and testing the multiple regression model, several assumptions were made concerning the geographical and unit characteristics. These assumptions are listed below.

1. Salaries increased four percent annually during the period 1979 to 1988. This is important because the most current tabulated figures for median household income by county is 1979 [Ref. 16]. In order to arrive at an income for 1988, the 1979 income figures were adjusted four percent annually.

2. Propensity to commute can be captured by using the percent of workers that commute to work outside of their residential county. The data for each county were obtained from the Bureau of the Census [Ref. 17].

3. The potential recruiting population for the reserves is adequately represented by the number of male category III-A's residing in the reserve center county.

4. Recruiting priorities for a reserve center are oriented toward filling the largest unit. Therefore, a reserve center can be categorized into one of four unit types (i.e., combat, combat support, combat service support, and medical) if fifty-one percent or more of the assigned personnel belong to that type of unit.

5. Regional differences in lifestyles are captured by dividing the county into nine census regions.

6. The percentage of civilian workers commuting outside of their residential county remains unchanged during the period 1986 to 1988. This is important because the most current data on the percentage of workers commuting outside of their residential county is 1986 [Ref. 17].

C. THE GENERAL MODEL

Multiple linear regression is used to describe a dependent variable by several independent explanatory variables. The general model is shown below [Ref. 10:p. 245]:

$$y_i = b_0 + \sum_{k=1}^{m} b_k x_{ik} + e_i \qquad \text{for } i = 1 \text{ to } n$$

where y_i are the n observed values of the response variable,

x_{ik} are the n values of the kth explanatory variable,

e_i are the normally distributed error terms,

and b_0, b_1, ..., b_k are the unknown regression coefficients.

In order to use multiple regression the following assumptions must be satisfied [Ref. 18:p. 412].

1. The relationship between the dependent and independent variables is linear.

2. The x values are observed without error. In addition, x values are known constants associated with the random variable y.

3. The y values are mutually independent random variables.

4. The variance of y is constant for all x values.

5. e_i is normally distributed with a zero mean and constant variance.

D. THESIS MODEL

The model for this thesis is shown below.

$$y_i = b_0 + \sum_{k=1}^{7} b_k x_{ik} + e_i \qquad \text{for } i = 1 \text{ to } 91$$

where x_1 = male category III-A population in the county

x_2 = percentage of males assigned to the reserve center

x_3 = annual median household income divided by the average annual military income at the reserve center

x_4 = unemployment rate in the reserve center county

x_5 = total number of male and female prior service personnel divided by the total number of personnel assigned to the reserve center

x_6 = percentage of workers that commute to work outside of their residential county

x_7 = the census region or the unit type.

The model assumptions are listed below.

1. The relationship between μ, σ, and θ and the selected independent variables is linear.
2. Values for each of the independent variables are known.
3. The values of m, s, and t are mutually independent random variables.
4. The variances of m, s, and t are constant for the seven explanatory variables.

5. The distributions of m, s, and t are normal.

The implicit null hypothesis is that the variable coefficients equal 0.

The model was run using the General Linear Model Procedure (Proc Glm) in SAS [Ref. 19]. Initially, it was tested with the first six explanatory variables. The seventh explanatory variable, census region or unit type, was not considered in order to determine if the model can be applied nationwide without regard to census region or unit type. Statistics from the analysis of variance tables for m, s, and t are summarized in Table 26 in Appendix N. To insure that the assumption of normality was valid, the residuals for m, s, and t were evaluated using quantile plots. Shown below in Figure 14 is the quantile plot with Kolmogorov-Smirnov bounds of the residuals for the t values.

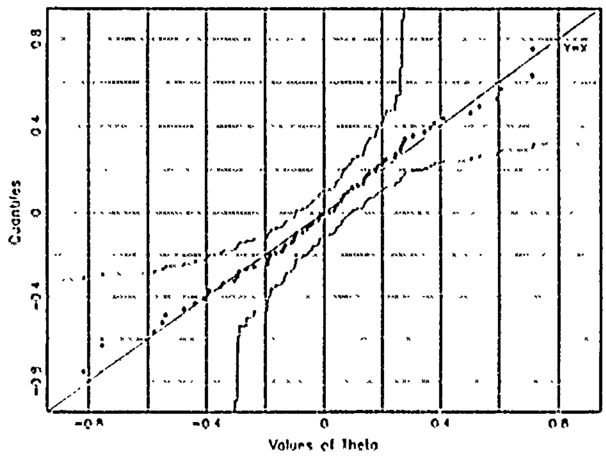

Figure 14

Plot of Theta Residuals

The residuals lie along the Y=X line and clearly fall inside the Kolmogorov-Smirnov bounds. Therefore, the assumption of normality is valid when the dependent variable is t. Similar quantile plots were evaluated for the m and s residuals but these failed the Kolmogorov-Smirnov test. The reasons for failure are the m and s estimates from the reserve center at Fond du Lac, Wisconsin. Referring to Table 24 in Appendix K, one observes a great disparity between the actual and estimated commuting distances for the 75th percentile. According to Table 4 in Appendix A, 88 out of 160 distance data points were adjusted because the reservists reside in the same zip code region as the reserve center. This inordinately high number of adjustments is the most likely cause for the outliers in the quantile plots for the m and s residuals. Because over 55 percent of the data points were modified, it is not unreasonable to drop the m and s values. After elimination, the residual plots pass the Kolmogorov-Smirnov test. Therefore, one can conclude that the assumption of normality is also valid when the dependent variables are m and s. Model results for m, s, and t are shown in Tables 27-29 in Appendix O.

The model was then tested including the nine census regions as the seventh explanatory variable. Census regions were selected as a test variable rather than the four geographical regions (i.e., northeast, midwest, south, and west) because census regions are commonly used to describe the United States. Model results for m, s, and t are shown in Tables 30-32 in Appendix O.

Final model testing included unit type as the seventh explanatory variable. Model results for m, s, and t are shown in Tables 33-35 in Appendix P.

E. MODEL RESULTS

1. The initial model pinpoints the variables that USAREC should carefully evaluate when assessing a geographical area. When applied nationwide, the full model R-Squared values for m, s, and t are between 0.1104 and 0.1619 (Table 26). These are not particularly high values, indicating that the model does not predict a significant portion of the dependent

variables' variance. For example, the model predicts at most 16.2% of the variance in the t values. One possible reason for the low R-Squared values is the fact that much of the data is estimated. For example, the commute data consists of Euclidean distances and not the actual distances reservists travel to their reserve centers. Another possible reason is that many important factors that impact on an individual's decision to commute may not have been captured in the data. For example, the road network in an area or the cost of fuel has not been captured in this model. However, upon examining Tables 27-29 one can identify those independent variables that have the greatest influence when predicting values of m, s, and t. One important statistic when evaluating the model is the last column (the probability that the variable is greater than the absolute value of the t statistic). Low probabilities indicate that the independent variable has a significant impact on the dependent variable. For example, in Table 27 in Appendix N the variable, propensity to commute, equals 0.0097. If our significance level is 0.05 (the value of $\alpha=0.05$) then propensity to commute is a significant variable. Conversely, the p-value for unemployment is greater than 0.05. Hence, unemployment is not a significant variable. Therefore, the significant variables are those that have a probability less than 0.05.

The parameter estimates in Tables 27-29 in Appendix O can be used to construct the multiple regression equations. For example, the regression equation for the dependent variable t from Table 27 is shown below.

$$t = -0.7518174 + 0.6258805\ x_2 + 0.0063553\ x_6$$

Equations for m and s can be constructed in a similar manner from Tables 28 and 29.

2. The second series of multiple regression equations were calculated using the nine census regions as a dummy variable. The analysis was conducted in SAS using Proc GLM commands with census regions as a class variable [Ref. 19:p. 434-506]. Using the value in the last column (the probability that the variable is greater than the absolute value of the t statistic), the mountain census region is a significant predicator for m and t when α equals 0.05. The variable s is best predicted by the mountain census region but it is not significant unless α is greater than 0.0532. Except for the mountain census region, it appears that there is no difference in s for nine census regions.

3. Unit type was used as a dummy variable in establishing the last three multiple regression equations. Again, the t statistic was used to determine if unit type was a significant variable. The lowest value was 0.4191 for variable "combat support" located in Table 33 in Appendix Q. Because all the t statistic values exceed 0.05, unit type, as measured in this model, does not appear to be not a significant variable for predicting values for m, s, and t.

F. CONCLUSIONS

The purpose of this thesis was to analyze individual commuting distances of enlisted U.S. Army Reservists and estimate a statistical relationship between commute behavior and different market variables. After analyzing commute distances from a sample of 91 reserve centers, it was determined that either the logistic or normal distribution with a suitable power transformation could accurately predict commuting distances. Various combinations of regional and unit characteristics were tested with the goal of developing a multiple regression equation to predict the parameters of the normal distribution. Although the R-Squared values were low, the model does identify civilian propensity to commute and percentage of male personnel in the unit as the two most influential independent variables.

Further research should focus on several independent variables not available in the 1988 data base established by the Defense Manpower Data Center. Four variables that would be invaluable are listed below.

1. Cost of fuel in the various census regions.
2. Individual commuting times.
3. Local road conditions.
4. Road access to each reserve center. For example, a reserve center located adjacent to an interstate might induce individuals to commute farther than one located off a secondary road.

Once these and other pertinent variables are included in the regression model the ability to assess a geographical region may be greatly enhanced.

APPENDIX A

TABLE 2

NORTHEASTERN UNITED STATES

Center Location	Zipcode	n	Missing	>100 Mi.	Same Zip
East Windsor, CT	06088	539	23	19	6
Middletown, CT	06457	135	3	7	15
West Hartford, CT	06110	575	28	25	5
Boston, MA	02210	883	38	52	11
Lawrence, MA	01843	288	24	7	10
Roslindale, MA	02131	255	14	9	9
Bangor, ME	04401	112	2	7	25
Dexter, ME	04930	118	9	3	4
Pedrickton, NJ	08067	474	21	9	6
Canadaigua, NY	14424	163	22	2	19
Canton, NY	13617	152	3	10	26
Elizabethtown, NY	12932	104	4	21	2
Glens Falls, NY	12801	123	3	4	31
Allison, PA	15601	304	19	7	52
Erie, PA	16504	355	17	17	17
Huntingdon, PA	16652	119	3	9	12
Philadelphia, PA	19154	319	14	8	7
Uniontown, PA	15401	350	20	5	67
TOTAL		5368	267	221	324

Percent of data used: 91

Percent unknown: 05

Percent greater than 100 miles: 04

Percent in the same zipcode: 06

TABLE 3

SOUTHERN UNITED STATES

Center Location	Zipcode	n	Missing	>100 Mi.	Same Zip
Decatur, AL	35601	156	2	5	18
Dothan, AL	36302	117	3	0	4
East Camden, AR	71701	259	8	8	62
El Dorado, AR	71730	187	3	10	61
Little Rock, AR	72204	403	10	28	49
West Memphis, AR	72301	120	6	19	17
Seaford, DE	19973	183	12	0	16
Tallahassee, FL	32304	365	11	49	108
Bardstown, KY	40004	121	0	3	17
Slidell, LA	70459	169	6	3	15
Hattiesburg, MS	39401	110	2	4	38
Jackson, MS	39209	509	5	22	91
Greensboro, NC	27409	290	5	19	4
Wilmington, NC	28401	236	1	8	51
Dewey, OK	74003	102	1	9	12
Greenwood, SC	29646	155	0	5	45
Carrollton, TX	75006	387	24	45	10
Ft. Bliss, TX	79906	152	7	5	3
Houston, TX	77054	1080	11	63	15

TABLE 3 (CONTINUED)

Center Location	Zipcode	n	Missing	>100 Mi	Same Zip
Paris, TX	75460	147	8	4	36
Victoria, TX	77901	115	3	5	39
Ft. Story, VA	23459	681	28	57	12
Richmond, VA	23220	573	24	43	21
Salem, VA	24153	320	12	16	20
Charleston, WV	25313	529	34	24	13
TOTAL		7466	226	454	777

Percent of data used: 91

Percent unknown: 03

Percent greater than 100 miles: 06

Percent in the same zipcode: 10

TABLE 4

MIDWESTERN UNITED STATES

Center Location	Zipcode	n	Missing	>100 Mi.	Same Zip
Ames, IA	50010	158	0	23	45
Council Bluffs, IA	51501	166	5	9	47
Decorah, IA	52101	196	2	3	30
Dubuque, IA	52001	144	2	14	59
Chicago, IL	60629	545	14	17	20
Granite City, IL	62040	384	6	17	22
Quincy, IL	62301	111	1	7	28
Springfield, IL	62703	102	1	9	13
Fort Wayne, IN	46809	522	8	19	13
Gary, IN	46404	208	0	5	16
Lake Station, IN	46405	307	4	9	0
Scottsburg, IN	47170	136	0	1	14
Bay City, MI	48706	306	4	9	35
Grand Rapids, MI	49503	508	7	27	26
Duluth, MN	55802	247	1	19	4
Fergus Falls, MN	56537	122	1	11	14
Columbia, MO	65201	256	6	22	56
Washington, MO	63640	111	1	3	20
Columbus, OH	43215	713	12	48	10

TABLE 4 (CONTINUED)

Center Location	Zipcode	n	Missing	>100 Mi	Same Zip
Toledo, OH	43606	495	4	17	20
Beloit, WI	53511	101	0	2	38
Fond du Lac, WI	54935	160	0	7	88
Wausau, WI	54401	225	3	20	43
TOTAL		6223	82	318	661

Percent of data used: 91

Percent unknown: 03

Percent greater than 100 miles: 06

Percent in the same zipcode: 10

TABLE 5

WESTERN UNITED STATES

Center Location	Zipcode	n	Missing	>100 Mi.	Same Zip
Tucson, AZ	85713	582	49	93	29
Bakersfield, CA	93301	191	22	20	5
Chico, CA	95926	124	10	14	27
Concord, CA	94519	140	2	2	4
Dublin, CA	94566	590	18	38	19
El Monte, CA	91733	639	56	41	14
Long Beach, CA	90822	184	17	16	0
Redding, CA	96003	112	3	10	13
San Pablo, CA	94806	431	70	30	17
Upland, CA	91786	136	7	10	1
Vallejo, CA	94589	113	3	2	7
Van Nuys, CA	91403	246	20	7	1
Aurora, CO	80011	459	43	53	21
Aurora, CO	80045	823	43	157	42
Denver, CO	80225	389	15	20	17
Hayden Lake, ID	83835	137	10	6	1
Great Falls, MT	59403	124	8	14	24
Helena, MT	59601	260	14	84	74
Las Cruces, NM	88001	188	3	17	59

TABLE 5 (CONTINUED)

Center Location	Zipcode	n	Missing	>100 Mi	Same Zip
Reno, NV	89502	241	21	45	21
Eugene, OR	97402	186	13	30	21
Logan, UT	84321	107	7	2	42
Ogden, UT	84407	403	26	22	31
Tacoma, WA	98404	316	21	34	14
Spokane, WA	99216	294	18	42	18
TOTAL		6223	82	318	661

Percent of data used: 91

Percent unknown: 03

Percent greater than 100 miles: 06

Percent in the same zipcode: 10

APPENDIX B

TABLE 6

NORTHEASTERN UNITED STATES

Center Location	Mileage Adjustment
East Windsor, CT	2.50
Middletown, CT	3.00
West Hartford,CT	3.25
Boston, MA	3.50
Lawrence, MA	3.50
Roslindale, MA	2.00
Bangor, ME	6.50
Dexter, ME	5.00
Pedrickton, NJ	3.00
Canadaigua, NY	5.50
Canton, NY	4.50
Elizabethtown, NY	6.50
Glens Falls, NY	5.00
Allison, PA	2.00
Erie, PA	5.50
Huntingdon, PA	5.50
Philadelphia, PA	4.50
Uniontown, PA	3.00

TABLE 7

SOUTHERN UNITED STATES

Center Location	Mileage Adjustment
Decatur, AL	6.00
Dothan, AL	4.50
East Camden, AR	6.00
El Dorado, AR	6.00
Little Rock, AR	5.00
West Memphis, AR	6.50
Seaford, DE	6.50
Tallahassee, FL	4.50
Bardstown, KY	6.00
Slidell, LA	5.50
Hattiesburg, MS	6.50
Jackson, MS	6.00
Greensboro, NC	5.00
Wilmington, NC	5.50
Dewey, OK	5.00
Greenwood, SC	6.50
Carrollton, TX	3.00
Ft. Bliss, TX	3.00
Houston, TX	12.50

TABLE 7 (CONTINUED)

Center Location	Mileage Adjustment
Paris, TX	5.00
Victoria, TX	5.00
Ft. Story, VA	5.00
Richmond, VA	6.00
Salem, VA	2.50
Charleston, WV	1.50

TABLE 8

MIDWESTERN UNITED STATES

Center Location	Mileage Adjustment
Ames, IA	5.00
Council Bluffs, IA	6.50
Decorah, IA	6.00
Dubuque, IA	5.00
Chicago, IL	6.00
Granite City, IL	4.50
Quincy, IL	6.00
Springfield, IL	7.50
Fort Wayne, IN	5.00
Gary, IN	2.50
Lake Station, IN	2.00
Scottsburg, IN	5.00
Bay City, MI	4.50
Grand Rapids, MI	4.00
Duluth, MN	3.00
Fergus Falls, MN	8.50
Columbia, MO	5.00
Washington, MO	6.50
Columbus, OH	5.00

TABLE 8 (CONTINUED)

Center Location	Mileage Adjustment
Toledo, OH	3.50
Beloit, WI	7.00
Fond du Lac, WI	4.00
Wausau, WI	6.00

TABLE 9

WESTERN UNITED STATES

Center Location	Mileage Adjustment
Tucson, AZ	3.00
Bakersfield, CA	2.50
Chico, CA	7.50
Concord, CA	4.50
Dublin, CA	5.00
El Monte, CA	1.50
Long Beach, CA	4.75
Redding, CA	5.00
San Pablo, CA	7.50
Upland, CA	2.00
Vallejo, CA	7.00
Van Nuys, CA	3.00
Aurora, CO	5.00
Aurora, CO	5.00
Denver, CO	3.50
Hayden Lake, ID	8.50
Great Falls, MT	6.00
Helena, MT	7.50
Las Cruces, NM	7.50

TABLE 9 (CONTINUED)

Center Location	Mileage Adjustment
Reno, NV	2.50
Eugene, OR	5.00
Logan, UT	5.00
Ogden, UT	4.00
Tacoma, WA	3.50
Spokane, WA	5.00

APPENDIX C

In the GAMS program for the logistic function, the equation LIKE was used if $\theta \neq 0$ and LIKE2 was used if $\theta = 0$. Listed below is the derivation of each equation.

Equation Like

$$L_\theta(\alpha, \beta) = \prod_{i=1}^{n} g(x_i)$$

Substituting in the equation for $g(x_i)$, one obtains:

$$L_\theta(\alpha, \beta) = \prod_{i=1}^{n} \frac{e^{-\frac{(x_i^\theta - \alpha)}{\beta}}}{\beta\left(1 + e^{-\frac{(x_i^\theta - \alpha)}{\beta}}\right)^2} \; \theta \, x^{\theta - 1}$$

The equation can be manipulated into the following form:

$$L_\theta(\alpha, \beta) = \frac{e^{-\frac{(\sum_{i=1}^{n} x_i^\theta - n\,\alpha)}{\beta}}}{\beta^n \prod_{i=1}^{n}\left(1 + e^{-\frac{(x_i^\theta - \alpha)}{\beta}}\right)^2} \; (\theta^n) \left(\prod_{i=1}^{n} x_i\right)^{\theta - 1}$$

By taking the natural logarithm, one obtains:

$$L_\theta(\alpha,\beta) = -\frac{(\sum_{i=1}^{n} x_i^\theta - n\alpha)}{\beta} + (n \ln \theta) + (\theta-1)\sum_{i=1}^{n} x_i$$

$$- n \ln \beta - 2\sum_{i=1}^{n} \ln \left(1 + e^{-\frac{(x_i^\theta-\alpha)}{\beta}}\right)$$

Equation Like2

$$L_\theta(\alpha,\beta) = \prod_{i=1}^{n} \frac{e^{-\frac{(\ln x_i^\theta-\alpha)}{\beta}}}{\beta\left(1 + e^{-\frac{(\ln x_i^\theta-\alpha)}{\beta}}\right)^2} \left(\frac{1}{x_i}\right)$$

$$L_\theta(\alpha,\beta) = \frac{e^{-\frac{(\sum_{i=1}^{n} \ln x_i^\theta - n\alpha)}{\beta}}}{(\beta^n) \prod_{i=1}^{n}\left(1 + e^{-\frac{(\ln x_i^\theta-\alpha)}{\beta}}\right)^2} \left(\frac{1}{\prod_{i=1}^{n} x_i}\right)$$

Taking the natural logarithm:

$$L_\theta(\alpha,\beta) = -\frac{(\sum_{i=1}^{n} \ln x_i^\theta - n\alpha)}{\beta} - n \ln \beta - 2\sum_{i=1}^{n} \ln \left(1 + e^{-\frac{(\ln x_i^\theta - \alpha)}{\beta}}\right) - \sum_{i=1}^{n} \ln x_i$$

GAMS numerically solves equations Like and Like2 for α, β, and θ (these estimates are referred to as a, b, and t in Chapter II).

APPENDIX D

TABLE 10

WESTERN UNITED STATES: LOGISTIC DISTRIBUTION

Center Location	Zipcode	a	b	t
Tucson, AZ	85713	2.062	0.335	-0.3203
Bakersfield, CA	93301	1.422	0.162	-0.1535
Chico, CA	95926	1.240	0.082	-0.1249
Concord, CA	94519	0.540	0.092	0.3647
Dublin, CA	94566	0.905	0.023	0.0673
El Monte, CA	91733	0.631	0.056	0.2173
Long Beach, CA	90822	0.832	0.036	0.1020
Redding, CA	96003	1.289	0.083	-0.1314
San Pablo, CA	94806	0.610	0.080	0.3013
Upland, CA	91786	0.748	0.053	0.1540
Vallejo, CA	94589	1.010	0.002	-0.0053
Van Nuys, CA	91403	0.837	0.033	0.1006
Aurora, CO	80011	1.606	0.184	-0.2258
Aurora, CO	80045	1.148	0.050	-0.0735
Denver, CO	80225	1.444	0.125	-0.1680
Hayden Lake, ID	83835	1.824	0.220	-0.3918
Great Falls, MT	59403	3.241	0.621	-0.4692
Helena, MT	59601	2.256	0.507	-0.4229
Las Cruces, NM	88001	2.479	0.496	-0.4481
Reno, NV	89502	4.329	0.998	-0.5384
Eugene, OR	97402	2.400	0.430	-0.4140
Logan, UT	84321	4.020	0.967	-0.5649
Ogden, UT	84407	1.018	0.005	-0.0088
Tacoma, WA	98404	1.502	0.124	-0.1740
Spokane, WA	99216	2.333	0.452	-0.3858

TABLE 11

MIDWESTERN UNITED STATES: LOGISTIC DISTRIBUTION

Center Location	Zipcode	a	b	t
Ames, IA	50010	1.094	0.034	-0.0452
Council Bluffs, IA	51501	3.055	0.635	-0.4813
Decorah, IA	52101	0.741	0.083	0.2317
Dubuque, IA	52001	2.639	0.612	-0.4219
Chicago, IL	60629	1.703	0.125	-0.2030
Granite City, IL	62040	3.125	0.483	-0.5208
Quincy, IL	62301	0.732	0.215	0.2152
Springfield, IL	62703	1.234	0.078	-0.0989
Fort Wayne, IN	46809	1.403	0.122	-0.1735
Gary, IN	46404	1.587	0.134	-0.1713
Lake Station, IN	46405	1.903	0.168	-0.2453
Scottsburg, IN	47170	0.469	0.088	0.4770
Bay City, MI	48706	0.800	0.054	0.1328
Grand Rapids, MI	49503	0.820	0.060	0.1122
Duluth, MN	55802	1.272	0.074	-0.1019
Fergus Falls, MN	56537	0.557	0.100	0.5511
Columbia, MO	65201	2.172	0.389	-0.3507
Columbus, OH	43215	1.430	0.149	-0.1564
Washington, MO	63640	0.954	0.014	0.0302
Toledo, OH	43606	1.652	0.214	-0.1933
Beloit, WI	53511	3.995	0.929	-0.6248
Fond du Lac, WI	54935	-2.670	0.500	0.0000
Wausau, WI	54401	1.034	0.011	-0.0199

TABLE 12

SOUTHERN UNITED STATES: LOGISTIC DISTRIBUTION

Center Location	Zipcode	a	b	t
Decatur, AL	35601	0.477	0.107	0.6003
Dothan, AL	36302	0.746	0.070	0.1469
East Camden, AR	71701	0.856	0.046	0.1000
El Dorado, AR	71730	0.842	0.053	0.1075
Little Rock, AR	72204	1.809	0.233	-0.2511
West Memphis, AR	72301	3.202	0.496	-0.5349
Seaford, DE	19973	0.677	0.062	0.2633
Tallahassee, FL	32304	1.003	0.001	-0.0015
Bardstown, KY	40004	0.612	0.080	0.3335
Slidell, LA	70459	0.406	0.092	0.7513
Hattiesburg, MS	39401	0.975	0.010	0.0152
Jackson, MS	39209	3.885	0.956	-0.5219
Greensboro, NC	27409	0.970	0.010	0.0192
Wilmington, NC	28401	2.675	0.480	-0.4495
Dewey, OK	74003	0.431	0.120	0.6938
Greenwood, SC	29646	1.245	0.070	-0.1179
Carrollton, TX	75006	0.686	0.062	0.2228
Ft. Bliss, TX	79906	3.608	0.691	-0.4593
Houston, TX	77054	1.169	0.037	-0.0729
Paris, TX	75460	0.718	0.097	0.2303
Victoris, TX	77901	0.740	0.090	0.1918
Ft. Story, VA	23459	1.261	0.049	-0.1250
Richmond, VA	23220	2.167	0.427	-0.3200
Salem, VA	24153	1.411	0.105	-0.1647
Charleston, WV	25313	0.898	0.025	0.0542

TABLE 13

NORTHEASTERN UNITED STATES: LOGISTIC DISTRIBUTION

Center Location	Zipcode	a	b	t
East Windsor, CT	06088	1.317	0.097	-0.1548
Middletown, CT	06457	0.962	0.014	0.0225
West Hartford, CT	06110	0.917	0.022	0.0417
Boston, MA	02210	1.475	0.120	-0.1515
Lawrence, MA	01843	0.602	0.069	0.2394
Roslindale, MA	02131	4.911	1.005	-0.4952
Bangor, ME	04401	1.616	0.215	-0.2557
Dexter, ME	04930	0.476	0.075	0.5834
Pedrickton, NJ	08067	0.681	0.059	0.2229
Canadaigua, NY	14424	0.646	0.075	0.2790
Canton, NY	13617	0.640	0.065	0.2390
Elizabethtown, NY	12932	0.837	0.041	0.1240
Glens Falls, NY	12801	1.628	0.173	-0.2282
Allison, PA	15601	0.602	0.072	0.2267
Erie, PA	16504	1.538	0.173	-0.1645
Huntingdon, PA	16652	1.109	0.020	-0.0541
Philadelphia, PA	19154	0.582	0.045	0.2494
Uniontown, PA	15401	0.652	0.064	0.1990

APPENDIX E

A GAMS program was used to numerically solve the MLE's (a, b, and t) using the logistic distribution for every reserve center. The input was formatted by using a fortran program. An example of the GAMS program is shown below. The reserve center depicted is located at Bakersfield, California.

PROGRAM

```
$OFFSYMLIST OFFSYMXREF

OPTIONS SOLPRINT = On, LIMCOL = 0, LIMROW = 0;

sets

$include file set;  (this file assigns a number to each commuter;
                     in this example, the numbers are 1 to 149)

scalar count

$include file count;  (this file gives a scalar for the number of
                       commuters to the reserve center; in this
                       example, count equals 149)

parameter X(I)

$include file out;  (this file is a listing of the commute
                     distances)

variable
     value (this is the value of the maximum likelihood function)
     al    (this is alpha)
     be    (this is beta)
     th;   (this is theta, the value of the power transformation)

positive variable be;

be.lo = 0.010;  (the range for beta is: 0.00 < beta ≤ 1.0)
be.up = 1.000;
al.lo = 0.010;  (the range for alpha is:  0.00 < alpha < 5.0)
al.up = 2.000;
th.lo = 0.010;  (the range for theta is:  0.00 < theta ≤ 1.0)
th.up = 1.000;
al.l  = 0.180;  (these four lines establish a starting point for
be.l  = 0.280;   the algorithm;  the "l" means level)
th.l  = 0.250;
value = 0.000;
```

```
equations
     LIKE2
     LIKE1;

LIKE1..  value =e=
     count*(al/be) - SUM(I,(X(I)**th))/be
   - count*LOG(be) - 2*SUM(I,LOG(1 + EXP(-((X(I)**th) - al)/be)))
   + count*(LOG(ABS(th)) + (th-1)*SUM(I,LOG(X(I)));

Model Stat /LIKE/;
Solve Stat using DNLP maximizing value;  (this directs the solver
                        to find the maximum value for the function
                        LIKE using the Discontinuous Non-Linear
                        Program algorithm)

Display  be.l, al.l, th.l, value.l;  (this displays the values for
                        alpha, beta, theta, and the function LIKE1)

be.lo =  0.006;  (the next series of lines sets new bounds and
be.up =  1.000;  starting values for alpha, beta, theta and value)
al.lo =  0.010;
al.up =  3.600;
th.lo = -1.000;  (the range for theta is: -1.00 ≤ theta < 0.00)
th.up = -0.001;
al.l  =  0.250;
be.l  =  0.180;
th.l  = -0.200;

Solve Stat using DNLP maximizing value;

Display be.l, al.l, tn.l, value.l;

LIKE2.. value =e=
     - count*LOG(be) - SUM(I,LOG(X(I)))/be + count*(al/be)
     - SUM(I,LOG(X(I)))
     - 2*SUM(I,LOG(1 + EXP(-(LOG(X(I))-al)/be)));

Model Stat2 /LIKE2/;

be.lo =  0.250;  (the next series of lines sets the bounds and
be.up =  1.000;   starting values for alpha, beta, and theta for
al.lo = -4.000;   the equation LIKE2)
al.up = -0.010;
be.l  =  0.280;
al.l  = -0.280;
value =  0.000;
```

```
Solve Stat2 using NLP maximizing value; (this directs the solver to
                                 use the Non-linear Programming algorithm)


Display be.l, al.l, value.l;  (the value of theta is zero so it
                                 is not being displayed)

be.lo    = 0.2500;  (the next series of lines sets new bounds and
be.up    = 2.0000;   starting values for alpha and beta)
al.lo    = 0.0005;
al.up    = 2.5000;
be.l     = 0.2800;
al.l     = 2.0000;
value.l  = 0.0000;

Solve Stat2 using NLP maximizing value;

Display be.l, al.l, value.l;
```

Output

	LOWER	LEVEL	UPPER	MARGINAL
EQU LIKE1	320.2040	320.2040	320.2040	1.0000

	LOWER	LEVEL	UPPER	MARGINAL
VAR Value	-INF	66.8657	+INF	.
VAR al	0.0100	0.9710	2.0000	EPS
VAR be	0.0100	0.0100	1.0000	-34.2779
VAR th	0.0100	0.0133	1.0000	EPS

```
REPORT SUMMARY:      0      NONOPT
                     0 INFEASIBLE
                     0  UNBOUNDED
                     0      ERRORS

VARIABLE be.l     =   0.010

VARIABLE al.l     =   0.971

VARIABLE th.l     =   0.013

VARIABLE value.l =  66.866
```

```
                 LOWER        LEVEL        UPPER        MARGINAL

EQU LIKE1      320.2040     320.2040     320.2040       1.0000

                 LOWER        LEVEL        UPPER        MARGINAL

VAR Value       -INF        68.9552       +INF            .
VAR al          0.0100       1.4218       3.6000         EPS
VAR be          0.0060       0.1623       1.0000         EPS
VAR th         -1.0000      -0.1535      -0.0100         EPS


REPORT SUMMARY:       0      NONOPT
                      0 INFEASIBLE
                      0  UNBOUNDED
                      0      ERRORS

VARIABLE be.l     =   0.162

VARIABLE al.l     =   1.422

VARIABLE th.l     =  -0.153

VARIABLE value.l =   68.955
```

```
                  LOWER         LEVEL         UPPER         MARGINAL

EQU LIKE2       320.2040      320.2040      320.2040        1.0000

                  LOWER         LEVEL         UPPER         MARGINAL

VAR value        -INF          67.1863        +INF             .
VAR al          -4.0000       -2.2136       -0.0100           EPS
VAR be           0.2500        0.7689        1.0000           EPS

REPORT SUMMARY:       0      NONOPT
                      0 INFEASIBLE
                      0  UNBOUNDED
                      0      ERRORS

VARIABLE be.l     =   0.769

VARIABLE al.l     =  -2.214

VARIABLE value.l =   67.186
```

	LOWER	LEVEL	UPPER	MARGINAL
EQU LIKE2	320.2040	320.2040	320.2040	1.0000

	LOWER	LEVEL	UPPER	MARGINAL
VAR Value	-INF	-34.5557	+INF	.
VAR al	0.0005	0.0005	2.5000	-53.5028
VAR be	0.2500	1.5248	2.0000	EPS

REPORT SUMMARY:		
	0	NONOPT
	0	INFEASIBLE
	0	UNBOUNDED
	0	ERRORS

VARIABLE be.l = 1.525

VARIABLE al.l = 5.000000E-4

VARIABLE value.l = -34.556

The optimal solution has the highest value for the maximum likelihood function. In this example, the highest value for the maximum likelihood function is 68.9552. The optimal parameters are:

1. alpha (a) = 1.4218
2. beta (b) = 0.1623
3. theta (t) = -0.1535

APPENDIX F

Shown below is the CDF of the logistic function.

$$G(x_i) = \begin{cases} \dfrac{1}{1 + e^{-\frac{(x_i{}^t - a)}{b}}} & if\ \ t \neq 0 \\[2ex] \dfrac{1}{1 + e^{-\frac{(\ln x_i - a)}{b}}} & if\ \ t = 0 \end{cases}$$

For a given percentile (call it P), one sets $G(x_i) = P$, obtaining:

$$P = \frac{1}{1 + e^{-\frac{(x_i{}^t - a)}{b}}} \qquad if\ t \neq 0$$

After appropriate mathematical manipulation, one can take the natural logarithm and obtain the following equation:

$$\frac{- x_i{}^t + a}{b} = \ln\left(\frac{1}{P} - 1\right)$$

Solving for x_i one obtains the following expression:

$$x_i = \begin{cases} \left(a - b\ (\ln\ (\frac{1}{P}-1))\right)^{\frac{1}{t}} & if\ \ t > 0 \\ \left(a - b\ (\ln\ (\frac{1}{1-P}-1))\right)^{\frac{1}{t}} & if\ \ t<0 \end{cases}$$

Similar calculations were done to obtain :

$$x_i = e^{(a\ -\ b\ (\ln\ (\frac{1}{P}-1)))} \qquad if\ t = 0$$

By substituting in values for a, b, t, and P the predicted commuting distance can be calculated.

APPENDIX G

TABLE 14

WESTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS LOGISTIC TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Tucson, AZ	582	9.7	3.4	17.2	19.3	62.8	41.4
Bakersfield, CA	191	7.1	10.1	34.1	24.1	77.4	65.9
Chico, CA	124	18.8	17.9	32.6	32.7	63.0	62.8
Concord, CA	140	19.2	18.5	30.2	29.6	41.5	44.2
Dublin, CA	590	23.7	22.7	31.7	34.2	60.4	50.9
El Monte, CA	639	12.5	12.0	18.1	18.4	26.9	27.3
Long Beach, CA	184	16.8	16.5	25.2	26.0	40.1	40.1
Redding, CA	112	10.4	14.5	34.1	25.3	49.6	46.3
San Pablo, CA	431	18.8	19.4	31.2	30.3	49.4	44.9
Upland, CA	136	16.6	15.2	22.9	24.7	37.4	38.8
Vallejo, CA	113	13.8	15.3	22.0	23.1	28.8	34.8
Van Nuys, CA	246	16.9	17.1	27.2	26.0	40.8	39.0
Aurora, CO	459	13.2	12.3	20.3	22.3	61.1	44.3
Aurora, CO	823	14.4	15.3	45.8	29.8	63.9	60.1
Denver, CO	389	11.0	11.2	19.6	20.4	55.4	39.4
Hayden Lake, ID	137	25.6	21.6	34.2	30.1	40.6	47.3
Great Falls, MT	124	6.0	8.2	19.2	13.5	57.6	26.2
Helena, MT	260	11.7	14.6	54.9	28.6	71.3	73.1
Las Cruces, NM	188	9.4	13.2	33.9	22.9	46.0	48.0

TABLE 14 (CONTINUED)

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Reno, NV	241	5.2	6.6	12.0	6.9	57.9	24.4
Eugene, OR	186	11.0	12.1	22.6	20.5	51.7	40.4
Logan, UT	107	8.9	8.5	20.9	14.7	37.5	32.2
Ogden, UT	403	10.0	13.2	35.6	24.4	42.4	45.2
Tacoma, WA	316	8.2	9.7	20.3	16.7	31.7	30.5
Spokane, WA	294	10.6	11.1	27.4	20.7	66.1	46.9
Mean Absolute Deviation		1.7 miles		5.2 miles		8.6 miles	

TABLE 15

SOUTHERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS LOGISTIC TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Decatur, AL	156	26.4	29.1	47.0	42.1	48.7	56.8
Dothan, AL	117	13.7	13.6	31.2	26.5	45.6	48.7
East Camden, AR	259	22.5	21.1	38.3	37.5	62.7	64.5
El Dorado, AR	187	26.6	20.2	41.7	37.6	69.7	67.4
Little Rock, AR	403	9.2	9.4	22.4	17.3	43.1	35.5
West Memphis, AR	120	12.1	11.4	16.5	16.1	28.2	24.7
Seaford, DE	183	21.2	22.7	32.1	32.7	43.4	45.6
Tallahassee, FL	365	15.6	13.6	17.7	28.2	37.9	58.6
Bardstown, KY	121	24.8	23.1	32.2	34.5	51.2	49.1
Slidell, LA	169	31.5	30.1	38.6	40.5	49.2	51.6
Hattiesburg, MS	110	20.4	18.9	63.0	39.5	72.2	81.9
Jackson, MS	509	6.0	7.4	22.3	13.6	52.8	33.0
Greensboro, NC	290	18.8	20.5	48.4	36.8	74.9	65.7
Wilmington, NC	236	10.1	11.2	19.6	18.3	44.4	34.2
Dewey, OK	102	33.3	29.7	43.9	43.7	51.7	59.2
Greenwood, SC	155	16.6	15.6	29.7	26.8	44.6	47.7
Carrollton, TX	387	18.5	18.4	25.2	28.2	47.4	41.5
Ft. Bliss, TX	152	6.4	6.1	9.1	10.2	37.2	20.1
Houston, TX	1080	12.4	11.7	18.9	19.1	30.3	31.6

TABLE 15 (CONTINUED)

Center Location	n	50th percentile		75th percentile		90th percentile	
		act.	est.	act.	est.	act.	est.
Paris, TX	147	29.7	23.7	42.5	43.3	77.2	73.4
Victoria, TX	115	24.2	20.8	52.9	40.0	58.8	71.5
Ft. Story, VA	681	15.5	15.6	21.2	22.2	31.8	31.9
Richmond, VA	573	6.7	8.9	25.3	19.1	71.7	52.5
Salem, VA	320	10.7	12.4	20.0	20.8	48.1	36.6
Charleston, WV	529	13.2	13.7	28.2	24.0	37.6	41.1
Mean Absolute Deviation		1.8 miles		4.5 miles		7.5 miles	

TABLE 16

MIDWESTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS LOGISTIC TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Ames, IA	158	13.0	13.7	35.4	29.6	72.1	65.5
Council Bluffs, IA	166	10.0	9.8	20.4	16.8	38.4	34.9
Decorah, IA	196	25.3	27.4	52.2	45.3	70.2	70.9
Dubuque, IA	144	8.2	10.0	29.2	20.1	69.1	54.3
Chicago, IL	545	7.6	7.3	11.2	11.0	18.9	17.3
Granite City, IL	384	10.7	11.2	14.8	16.0	28.6	24.9
Quincy, IL	111	24.0	23.5	42.9	40.5	67.4	66.0
Springfield, IL	102	13.5	11.9	32.7	24.7	58.7	54.1
Fort Wayne, IN	522	13.2	14.2	34.4	25.3	48.5	48.2
Gary, IN	208	6.2	7.0	15.1	12.4	21.5	23.3
Lake Station, IN	307	6.0	7.3	11.8	11.0	18.0	17.5
Scottsburg, IN	136	23.6	20.5	33.0	30.3	37.8	42.2
Bay City, MI	306	19.4	18.6	35.4	31.9	48.9	52.8
Grand Rapids, MI	508	19.9	17.1	37.6	34.0	57.4	64.4
Duluth, MN	247	10.4	9.4	16.3	18.0	46.7	36.1
Fergus Falls, MN	122	34.4	34.6	51.0	47.9	56.0	63.2
Columbia, MO	256	5.5	11.0	28.4	20.5	47.4	45.6
Washington, MO	111	21.2	21.0	46.6	35.7	58.3	60.1
Columbus, OH	713	8.5	10.2	35.8	22.1	58.5	53.6

TABLE 16 (CONTINUED)

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Toledo, OH	495	6.5	7.5	26.1	16.5	45.7	42.1
Beloit, WI	101	10.4	10.9	22.1	17.5	43.5	34.2
Fond du Lac, WI	160	4.0	6.9	20.6	12.0	28.1	20.8
Wausau, WI	225	18.3	18.6	34.0	33.6	56.6	61.2
Mean Absolute Deviation		1.3 miles		5.2 miles		4.6 miles	

TABLE 17

NORTHEASTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS LOGISTIC TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
East Windsor, CT	539	16.0	16.9	30.3	29.1	66.3	52.8
Middletown, CT	135	18.4	17.9	31.7	36.2	84.5	72.4
West Hartford,CT	575	12.6	12.5	27.5	23.4	41.7	42.9
Boston, MA	883	7.5	7.7	16.5	14.3	35.3	28.2
Lawrence, MA	288	13.1	12.0	18.7	19.7	28.8	30.7
Roslindale, MA	255	3.9	4.0	7.7	6.7	16.0	13.4
Bangor, ME	112	15.0	15.3	34.6	28.4	61.0	59.2
Dexter, ME	118	29.3	28.0	34.3	36.8	45.6	46.6
Pedrickton, NJ	474	19.3	17.8	25.1	26.8	38.2	39.0
Canadaigua, NY	163	25.2	20.9	29.7	32.1	40.9	47.2
Canton, NY	152	16.9	15.5	25.6	24.1	31.2	35.9
Elizabethtown, NY	104	23.7	23.8	35.2	36.3	51.2	54.3
Glens Falls, NY	123	10.2	11.8	26.0	20.4	36.9	37.9
Allison, PA	304	11.8	10.7	17.6	18.4	29.0	29.8
Erie, PA	355	6.0	7.3	22.4	16.3	43.9	41.0
Huntingdon, PA	119	13.6	14.8	21.5	21.4	30.2	31.2
Philadelphia, PA	319	11.1	11.4	15.6	15.8	19.6	21.4
Uniontown, PA	350	13.4	11.7	20.6	19.5	30.6	31.1
Mean Absolute Deviation		1.1 miles		2.4 miles		3.6 miles	

APPENDIX H

A GAMS program was used to numerically solve the MLE's (m,s, and t) using the normal distribution for every reserve center. The input was formatted by using a fortran program. An example of the GAMS program is shown below. The reserve center depicted is located at Bakersfield, California.

PROGRAM

```
$OFFSYMLIST OFFSYMXREF

OPTIONS SOLPRINT = On, LIMCOL = 0, LIMROW = 0;

sets

$include file set; (this file assigns a number to each commuter;
                    in this example, the numbers are from 1 to 149)

scalar count

$include file count; (this file gives a scalar for the number of
                      commuters to the reserve center; in this
                      example, count equals 149)

parameter X(I)

$include file out; (this file is a listing of the commute
                    distances)

variable
     value (this is the value of the maximum likelihood function)
     mu
     sigma
     th;   (this is theta, the value of the power transformation)

positive variable sigma;

th.lo   = 0.001; (these 5 lines establish bounds and set starting
th.up   = 1.000;  values for theta, mu, and sigma; sigma can only
mu.l     = 1.000;  be a positve value; mu can take on any positve
sigma.l = 5.000;  or negative value)
th.l    = 0.250;
value.l = 0.000; (initial value of the function is equal to zero)

equation
     NORMAL;
```

```
NORMAL..  value =e=
     ((-0.5*count)*LOG(2*3.1459)) - (count)*LOG(sigma)
   - (0.5/(sigma**2))*(SUM(I,(ABS((((X(I)**th)-1)/th)-mu)**2)))
   + (th-1)*SUM(I,LOG(X(I)));

Model Stat /NORMAL/;
Solve Stat using DNLP maximizing value;  (this directs the solver
                        to find the maximum value for the function
                        NORMAL using the Discontinuous Non-Linear
                        Program algorithm)

Display  mu.l, sigma.l, th.l, value.l;  (this displays the values
                     for mu, sigma, theta, and the function NORMAL)

th.lo    = -0.950;  (the value of theta is modified so that the
th.up    = -0.001;   function is evaluated from the left side
mu.l     =  1.000;   of zero)
sigma.l =  5.000;
th.l     = -0.200;
value.l =  0.000;

Solve Stat using DNLP maximizing value;

Display mu.l, sigma.l, th.l, value.l;

mu.l     = 3.000;
sigma.l = 5.000;
value.l = 0.000;

equation
     NORMAL2;

NORMAL2.. value =e=
     ((-0.5*count)*LOG(2*3.1459)) - (count)*LOG(sigma)
   - ((0.5/(sigma**2))*(SUM(I,(ABS(LOG(X(I))-mu)**2))))
   - SUM(I,LOG(X(I)));

Model Stat2 /NORMAL2/;

Solve Stat2 using DNLP maximizing value;

Display mu.l, sigma.l, value.l;  (the value of theta equals zero)
```

Output

	LOWER	LEVEL	UPPER	MARGINAL
EQU NORMAL	183.1801	183.1801	183.1801	1.0000

	LOWER	LEVEL	UPPER	MARGINAL
VAR Value	-INF	74.1694	+INF	.
VAR mu	-INF	-2.1459	+INF	EPS
VAR sigma	.	1.2579	+INF	EPS
VAR th	0.0010	0.0010	1.0000	-24.8407

```
REPORT SUMMARY:       0      NONOPT
                      0 INFEASIBLE
                      0  UNBOUNDED
                      0      ERRORS

VARIABLE mu.1    =  -2.146

VARIABLE sigma.1 =   1.258

VARIABLE th.1    =   0.001

VARIABLE value.1 =  74.169
```

	LOWER	LEVEL	UPPER	MARGINAL
EQU NORMAL	183.1801	183.1801	183.1801	1.0000

	LOWER	LEVEL	UPPER	MARGINAL
VAR Value	-INF	75.9896	+INF	.
VAR mu	-INF	-2.6778	+INF	EPS
VAR sigma	.	1.7013	+INF	EPS
VAR th	-0.9500	-0.1452	-0.0010	EPS

```
REPORT SUMMARY:       0      NONOPT
                      0 INFEASIBLE
                      0  UNBOUNDED
                      0      ERRORS

VARIABLE mu.1    =  -2.678

VARIABLE sigma.1 =   1.701

VARIABLE th.1    =  -0.145

VARIABLE value.1 =  75.990
```

	LOWER	LEVEL	UPPER	MARGINAL
EQU NORMAL2	183.1801	183.1801	183.1801	1.0000

	LOWER	LEVEL	UPPER	MARGINAL
VAR value	-INF	74.1942	+INF	.
VAR mu	-INF	-2.1490	+INF	EPS
VAR sigma	.	1.2604	+INF	EPS

```
REPORT SUMMARY:        0     NONOPT
                       0 INFEASIBLE
                       0  UNBOUNDED
                       0     ERRORS

VARIABLE mu.l     =  -2.149

VARIABLE sigma.l  =   1.260

VARIABLE value.l  =  74.194
```

The optimal solution has the highest value for the maximum likelihood function. In this example, the highest value for the maximum likelihood function is 75.9896. The optimal parameters are:

1. mu (m) = -2.6778
2. sigma (s) = 1.7013
3. theta (t) = -0.1452

APPENDIX I

TABLE 18

SOUTHERN UNITED STATES: NORMAL DISTRIBUTION

Center Location	Zipcode	m	s	t
Decatur, AL	35601	-0.859	0.227	0.6430
Dothan, AL	36302	-1.746	0.790	0.1455
East Camden, AR	71701	-1.488	0.786	0.0746
El Dorado, AR	71730	-1.556	0.862	0.0592
Little Rock, AR	72204	-3.101	1.515	-0.2320
West Memphis, AR	72301	-4.457	1.791	-0.5959
Seaford, DE	19973	-1.273	0.442	0.2264
Tallahassee, FL	32304	-2.378	1.029	-0.1000
Bardstown, KY	40004	-1.226	0.457	0.2745
Slidell, LA	70459	-0.814	0.231	0.7178
Hattiesburg, MS	39401	-1.600	1.017	0.0353
Jackson, MS	39209	-5.666	3.246	-0.5512
Greensboro, NC	27409	-1.526	0.866	0.0461
Wilmington, NC	28401	-3.941	1.949	-0.4976
Dewey, OK	74003	-0.839	0.289	0.6968
Greenwood, SC	29646	-2.210	1.060	-0.1742
Carrollton, TX	75006	-1.440	0.529	0.2036
Ft. Bliss, TX	79906	-6.036	2.876	-0.4968
Houston, TX	77054	-2.397	0.950	-0.0984
Paris, TX	75460	-1.264	0.708	0.2127
Victoria, TX	77901	-1.397	0.774	0.1760
Ft. Story, VA	23459	-2.187	0.776	-0.1648
Richmond, VA	23220	-3.568	2.190	-0.3207
Salem, VA	24153	-2.287	0.998	-0.1002
Charleston, WV	25313	-1.821	0.730	0.0849

TABLE 19

WESTERN UNITED STATES: NORMAL DISTRIBUTION

Center Location	Zipcode	m	s	t
Tucson, AZ	85713	-3.306	1.804	-0.3259
Bakersfield, CA	93301	-2.678	1.701	-0.1452
Chico, CA	95926	-2.003	1.144	-0.1639
Concord, CA	94519	-1.312	0.463	0.3310
Dublin, CA	94566	-1.404	0.590	0.0712
El Monte, CA	91733	-1.692	0.452	0.2204
Long Beach, CA	90822	-1.671	0.662	0.0879
Redding, CA	96003	-2.066	0.987	-0.0853
San Pablo, CA	94806	-1.275	0.439	0.3238
Upland, CA	91786	-1.721	0.650	0.1125
Vallejo, CA	94589	-2.165	0.713	-0.0896
Van Nuys, CA	91403	-1.628	0.563	0.0980
Aurora, CO	80011	-2.682	1.392	-0.2316
Aurora, CO	80045	-1.954	1.080	-0.0560
Denver, CO	80225	-2.663	1.297	-0.1760
Hayden Lake, ID	83835	-2.250	1.018	-0.4708
Great Falls, MT	59403	-3.290	1.614	-0.2581
Helena, MT	59601	-2.965	1.976	-0.4470
Las Cruces, NM	88001	-3.328	1.833	-0.4724
Reno, NV	89502	-6.144	3.278	-0.5450
Eugene, OR	97402	-3.518	1.879	-0.4468
Logan, UT	84321	-5.838	3.198	-0.6331
Ogden, UT	84407	-1.980	0.923	-0.0028
Tacoma, WA	98404	-2.833	1.175	-0.1651
Spokane, WA	99216	-3.117	1.784	-0.3118

TABLE 20

MIDWESTERN UNITED STATES: NORMAL DISTRIBUTION

Center Location	Zipcode	m	s	t
Ames, IA	50010	-1.986	1.164	-0.0174
Council Bluffs, IA	51501	-6.819	3.764	-0.8476
Decorah, IA	52101	-1.010	0.568	0.2611
Dubuque, IA	52001	-4.070	2.557	-0.4700
Chicago, IL	60629	-3.650	1.192	-0.2370
Granite City, IL	62040	-4.016	1.661	-0.5120
Quincy, IL	62301	-1.289	0.662	0.1910
Springfield, IL	62703	-2.403	1.325	-0.1145
Fort Wayne, IN	46809	-2.248	1.121	-0.1505
Gary, IN	46404	-4.189	1.811	-0.2700
Lake Station, IN	46405	-3.406	1.092	-0.1865
Scottsburg, IN	47170	-1.128	0.310	0.4713
Bay City, MI	48706	-1.550	0.706	0.1119
Grand Rapids, MI	49503	-1.613	0.855	0.1155
Duluth, MN	55802	-2.792	1.338	-0.1329
Fergus Falls, MN	56537	-0.823	0.321	0.5311
Columbia, MO	65201	-5.086	3.094	-0.5545
Washington, MO	63640	-1.519	0.780	0.0290
Columbus, OH	43215	-2.662	1.529	-0.1432
Toledo, OH	43606	-3.297	1.779	-0.1869
Beloit, WI	52511	-4.661	2.479	-0.6222
Fond du Lac, WI	54935	-12.345	7.044	-0.8806
Wausau, WI	54401	-1.754	0.915	-0.0513

TABLE 21

NORTHEASTERN UNITED STATES: NORMAL DISTRIBUTION

Center Location	Zipcode	m	s	t
East Windsor, CT	06088	-1.951	1.013	-0.1171
Middletown, CT	06457	-1.668	0.999	0.0302
West Hartford, CT	06110	-1.999	0.892	0.0369
Boston, MA	02210	-2.984	1.270	-0.1231
Lawrence, MA	01843	-1.733	0.552	0.2052
Roslindale, MA	02131	-8.718	3.983	-0.5463
Bangor, ME	04401	-2.511	1.460	-0.2993
Dexter, ME	04930	-0.923	0.251	0.5476
Pedrickton, NJ	08067	-1.478	0.499	0.1910
Canadaigua, NY	14424	-1.367	0.522	0.2047
Canton, NY	13617	-1.641	0.539	0.1579
Elizabethtown, NY	12932	-1.324	0.589	0.1197
Glens Falls, NY	12801	-2.869	1.336	-0.2648
Allison, PA	15601	-1.914	0.648	0.1601
Erie, PA	16504	-3.283	1.754	-0.1727
Huntingdon, PA	16652	-2.166	0.763	-0.1199
Philadelphia, PA	19154	-1.857	0.418	0.1572
Uniontown, PA	15401	-1.838	0.586	0.1615

APPENDIX J

Shown below is the normal CDF.

$$P = \int_{-\infty}^{v_i} \frac{1}{\sqrt{2\pi}\, s} e^{\left(\frac{-(v_i-m)}{s}\right)^2} dv_i$$

where P is the percentile commute distance.

A statistical package called GRAFSTAT has a function labelled NORICDF. By entering the values for m,s, and the desired percentile (0.5 for the 50th percentile commute distance, 0.75 for the 75th percentile, etc.) into the function NORICDF, one can obtain a value for v_i .

As discussed in Chapter III, the transformation selected is:

$$v_i = \frac{x_i^{\,t}-1}{t}$$

In order to obtain the estimated commute distance, the transformation must be solved for x_i . Therefore, the equation can be simplied to:

$$x_i = (tv_i+1)^{\frac{1}{t}}$$

The values of v_i and t are substituted into the above equation to obtain the estimated commute distance.

APPENDIX K

TABLE 22

WESTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS NORMAL TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Tucson, AZ	582	9.7	10.6	17.2	20.3	62.8	42.3
Bakersfield, CA	191	7.1	10.4	34.1	25.1	77.4	61.8
Chico, CA	124	18.8	17.7	32.6	32.6	63.0	59.8
Concord, CA	140	19.2	17.9	30.2	29.6	41.5	44.0
Dublin, CA	590	23.7	22.8	31.7	35.2	60.4	51.5
El Monte, CA	639	12.5	12.0	18.1	19.1	26.9	27.9
Long Beach, CA	184	16.8	16.4	25.2	27.4	40.1	42.6
Redding, CA	112	10.4	14.9	34.1	26.6	49.6	46.1
San Pablo, CA	431	18.8	19.3	31.2	30.8	49.4	44.5
Upland, CA	136	16.6	14.8	22.9	25.0	37.4	39.2
Vallejo, CA	113	13.8	13.8	22.0	20.8	28.8	30.6
Van Nuys, CA	246	16.9	17.0	27.2	25.3	40.8	38.7
Aurora, CO	459	13.2	12.4	20.3	23.1	61.1	44.2
Aurora, CO	823	14.4	15.7	45.8	30.6	63.9	57.1
Denver, CO	389	11.0	11.3	19.6	21.1	55.4	39.8
Hayden Lake, ID	137	25.6	21.6	34.2	31.0	40.6	45.8
Great Falls, MT	124	6.0	9.2	19.2	17.5	57.6	34.6
Helena, MT	260	11.7	15.1	54.9	29.2	71.3	67.4
Las Cruces, NM	188	9.4	13.5	33.9	23.3	46.0	44.7

TABLE 22 (CONTINUED)

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Reno, NV	241	5.2	6.7	12.0	12.2	57.9	26.6
Eugene, OR	186	11.0	12.1	22.6	21.1	51.7	40.6
Logan, UT	107	8.9	8.7	20.9	15.0	37.5	31.0
Ogden, UT	403	10.0	13.9	35.6	25.8	42.4	45.1
Tacoma, WA	316	8.2	9.8	20.3	17.2	31.7	30.1
Spokane, WA	294	10.6	11.3	27.4	22.3	66.1	47.8
Mean Absolute Deviation		1.7 miles		4.7 miles		8.5 miles	

TABLE 23

SOUTHERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS NORMAL TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Decatur, AL	156	26.4	28.7	47.0	41.5	48.7	79.2
Dothan, AL	117	13.7	13.9	31.2	26.3	45.6	46.0
East Camden, AR	259	22.5	20.7	38.3	37.0	62.7	61.3
El Dorado, AR	187	26.6	19.6	41.7	36.6	69.7	63.3
Little Rock, AR	403	9.2	9.7	22.4	18.3	43.1	35.8
West Memphis, AR	120	12.1	11.4	16.5	16.4	28.2	24.9
Seaford, DE	183	21.2	22.3	32.1	33.2	43.4	46.3
Tallahassee, FL	365	15.6	11.8	17.7	21.1	37.9	36.5
Bardstown, KY	121	24.8	22.4	32.2	34.7	51.2	49.5
Slidell, LA	169	31.5	29.4	38.6	41.0	49.2	52.3
Hattiesburg, MS	110	20.4	19.3	63.0	39.5	72.2	74.2
Jackson, MS	509	6.0	7.7	22.3	14.3	52.8	33.4
Greensboro, NC	290	18.8	20.6	48.4	38.2	74.9	65.7
Wilmington, NC	236	10.1	11.3	19.6	18.6	44.4	33.7
Dewey, OK	102	33.3	28.3	43.9	42.5	51.7	56.7
Greenwood, SC	155	16.6	15.4	29.7	26.5	44.6	45.2
Carrollton, TX	387	18.5	18.2	25.2	29.4	47.4	43.7
Ft. Bliss, TX	152	6.4	6.1	9.1	10.7	37.2	21.1
Houston, TX	1080	12.4	11.6	18.9	19.8	30.3	32.9

TABLE 23 (CONTINUED)

Center Location	n	50th percentile		75th percentile		90th percentile	
		act.	est.	act.	est.	act.	est.
Paris, TX	147	29.7	22.9	42.5	42.3	77.2	69.0
Victoria, TX	115	24.2	20.1	52.9	38.7	58.8	65.7
Ft. Story, VA	681	15.5	15.4	21.2	23.0	31.8	33.7
Richmond, VA	573	6.7	9.3	25.3	20.2	71.7	50.6
Salem, VA	320	10.7	12.8	20.0	22.4	48.1	38.3
Charleston, WV	529	13.2	13.8	28.2	24.4	37.6	39.8
Mean Absolute Deviation		2.1 miles		4.5 miles		7.1 miles	

TABLE 24

MIDWESTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS NORMAL TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
Ames, IA	158	13.0	14.2	35.4	30.5	72.1	61.1
Council Bluffs, IA	166	10.0	10.5	20.4	16.4	38.4	31.1
Decorah, IA	196	25.3	31.0	52.2	50.4	70.2	74.6
Dubuque, IA	144	8.2	10.3	29.2	20.6	69.1	51.0
Chicago, IL	545	7.6	7.2	11.2	11.4	18.9	17.9
Granite City, IL	384	10.7	11.3	14.8	16.9	28.6	26.7
Quincy, IL	111	24.0	22.8	42.9	39.9	67.4	63.1
Springfield, IL	102	13.5	12.0	32.7	24.8	58.7	50.8
Fort Wayne, IN	522	13.2	14.4	34.4	26.0	48.5	46.5
Gary, IN	208	6.2	6.1	15.1	11.3	21.5	22.0
Lake Station, IN	307	6.0	7.2	11.8	11.5	18.0	18.2
Scottsburg, IN	136	23.6	20.0	33.0	30.0	37.8	40.8
Bay City, MI	306	19.4	18.2	35.4	31.9	48.9	51.2
Grand Rapids, MI	508	19.9	16.8	37.6	33.2	57.4	58.8
Duluth, MN	247	10.4	9.3	16.3	18.5	46.7	36.6
Fergus Falls, MN	122	34.4	33.9	51.0	48.1	56.0	62.9
Columbia, MO	256	5.5	8.9	28.4	17.1	47.4	41.8
Washington, MO	111	21.2	21.2	46.6	36.5	58.3	59.3
Columbus, OH	713	8.5	10.5	35.8	23.1	58.5	55.1

TABLE 24 (CONTINUED)

Center Location	n	50th percentile		75th percentile		90th percentile	
		act.	est.	act.	est.	act.	est.
Toledo, OH	495	6.5	7.7	26.1	17.0	45.7	39.4
Beloit, WI	101	10.4	11.2	22.1	18.5	43.5	35.0
Fond du Lac, WI	160	4.0	6.0	20.6	9.9	28.1	21.2
Wausau, WI	225	18.3	18.7	34.0	33.1	56.6	56.4
Mean Absolute Deviation		1.5 miles		5.2 miles		5.0 miles	

TABLE 25

NORTHEASTERN UNITED STATES: A COMPARISON OF ACTUAL VERSUS NORMAL TRANSFORMATION COMMUTING DISTANCES

Center Location	n	50th Percentile		75th Percentile		90th Percentile	
		act.	est.	act.	est.	act.	est.
East Windsor, CT	539	16.0	17.3	30.3	30.7	66.3	53.3
Middletown, CT	135	18.4	18.1	31.7	36.4	84.5	67.7
West Hartford, CT	575	12.6	12.5	27.5	23.8	41.7	41.9
Boston, MA	883	7.5	7.9	16.5	15.1	35.3	28.5
Lawrence, MA	288	13.1	11.8	18.7	20.3	28.8	31.6
Roslindale, MA	255	3.9	4.1	7.7	6.9	16.0	13.6
Bangor, ME	112	15.0	15.4	34.6	28.4	61.0	55.7
Dexter, ME	118	29.3	27.6	34.3	37.8	45.6	48.2
Pedrickton, NJ	474	19.3	17.6	25.1	27.6	38.2	40.1
Canadaigua, NY	163	25.2	20.1	29.7	32.1	40.9	47.1
Canton, NY	152	16.9	15.0	25.6	24.0	31.2	35.7
Elizabethtown, NY	104	23.7	23.7	35.2	37.4	51.2	55.5
Glens Falls, NY	123	10.2	11.8	26.0	20.5	36.9	36.5
Allison, PA	304	11.8	10.2	17.6	18.5	29.0	30.4
Erie, PA	355	6.0	7.4	22.4	16.7	43.9	38.6
Huntingdon, PA	119	13.6	14.6	21.5	22.2	30.2	32.9
Philadelphia, PA	319	11.1	11.1	15.6	16.4	19.6	22.7
Uniontown, PA	350	13.4	11.3	20.6	19.3	30.6	30.3
Mean Absolute Deviation		1.2 miles		2.6 miles		4.4 miles	

APPENDIX L

Listed below are the nine census regions that are used in the multiple regression model. Also shown are the states that comprise each region.

Region	States
1. Pacific	Washington, Oregon, California
2. Mountain	Montana, Idaho, Nevada, Utah, Arizona, Colorado, Wyoming, New Mexico
3. West North Central	North Dakota, South Dakota, Iowa, Nebraska, Kansas, Minnesota, Missouri
4. West South Central	Texas, Oklahoma, Arkansas, Louisiana
5. East North Central	Wisconsin, Illinois, Michigan, Indiana, Ohio
6. East South Central	Kentucky, Tennessee, Mississippi, Alabama
7. New England	Maine, Vermont, Massachusetts, Connecticut, Rhode Island, New Hampshire
8. Mid-Atlantic	New York, Pennslyvania, New Jersey
9. South Atlantic	Maryland, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Florida, Delaware

APPENDIX M

Listed below are the types of units located at the reserve centers and the category to which they are assigned.

Category	**Unit Type**
1. Combat	Airborne, Armour, Field Artillery, Infantry, Special Forces
2. Combat Support	Headquarters, Air Defense, Aviation, Chemical, Engineers, Logistical Command, Military Intelligence, Military Police, Signal
3. Combat Service Support	Adjutant General, Civil Affairs, Finance, Judge Advocate General, Ordinance, Psychological Operations, Quartermaster, Transportation, Training, Schools
4. Medical	Medical

APPENDIX N

TABLE 26

SUMMARY STATISTICS FROM THE MULTIPLE REGRESSION ANALYSIS OF VARIANCE TABLES

	m	s	t
R-Square	0.1104	0.1360	0.1614
Model Sum of Squares	44.9493	12.6885	1.5417
Error Sum of Squares	362.2523	80.6151	7.9743
Total Sum of Squares	407.2016	93.3035	9.5209
Model Mean Square	7.4916	2.1147	0.2569
Error Mean Square	4.3125	0.9597	0.0950
F Value	1.7400	2.2000	2.7000
Prob. > F	0.1223	0.0505	0.0189

APPENDIX O

TABLE 27

PARAMETER ESTIMATES FROM THE
MULTIPLE REGRESSION MODEL FOR T

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	-0.7518174	0.3471955	0.0332
Male III-A Population	9.89541E-07	0.0000006	0.0916
Percentage of Males at the Reserve Center	0.6258805	0.2877690	0.0324
Civilian Wages/Military Wages	0.0047917	0.0165479	0.7729
Unemployment	0.0123455	0.0178909	0.4921
Percentage of Prior Service Personnel	-0.3333634	0.3445064	0.3360
Propensity to Commute	0.0063553	0.0024008	0.0097

TABLE 28

PARAMETER ESTIMATES FROM THE
MULTIPLE REGRESSION MODEL FOR M

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	-2.7087299	2.3393631	0.2502
Male III-A Population	4.48160E-06	0.0000039	0.2547
Percentage of Males at the Reserve Center	3.6032658	1.9389543	0.0666
Civilian Wages/Military Wages	-0.2350680	0.1114978	0.0380
Unemployment	-0.0343118	0.1205471	0.7766
Percentage of Prior Service Personnel	-1.7935977	2.3212443	0.4419
Propensity to Commute	0.0291351	0.0161764	0.0753

TABLE 29

PARAMETER ESTIMATES FROM THE
MULTIPLE REGRESSION MODEL FOR S

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	3.2768767	1.1035709	0.0039
Male III-A Population	-3.96858E-06	0.0000018	0.0342
Percentage of Males at the Reserve Center	-1.5044185	0.9146821	0.1038
Civilian Wages/Military Wages	0.0213407	0.0525979	0.6860
Unemployment	0.0644819	0.0568669	0.2601
Percentage of Prior Service Personnel	-0.5948843	1.0950236	0.5884
Propensity to Commute	-0.0184917	0.0076310	0.0175

APPENDIX P

TABLE 30

PARAMETER SUM OF SQUARES FROM THE MULTIPLE REGRESSION MODEL FOR S INCLUDING THE NINE CENSUS REGIONS

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	3.6295838	0.0000021	0.0962
Male III-A Population	-3.54941E-06	1.2976608	0.0065
Percentage of Males at the Reserve Center	-1.8115175	0.9680104	0.0651
Civilian Wages/Military Wages	-0.0104021	0.0583026	0.8589
Unemployment	-0.0698944	0.0721239	0.3356
Percentage of Prior Service Personnel	-0.8899148	1.4312514	0.5360
Propensity to Commute	-0.0167935	0.0081646	0.0431
Census Regions			
East North Central	0.6406002	0.4299761	0.1404
East South Central	0.1964427	0.5398399	0.7169
Mid-Atlantic	0.0473899	0.4648334	0.9191
Mountain	0.8647664	0.4403284	0.0532
New England	0.1205138	0.5495573	0.8270
Pacific	0.2712929	0.5105324	0.5967
South Atlantic	-0.0913791	0.5097881	0.8582
West North Central	0.5724342	0.5009444	0.2567
West South Central	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.2285
Model S.S. = 21.3213
Error S.S. = 71.9823
F Value = 1.61
Prob. > F = 0.0962

TABLE 31

PARAMETER SUM OF SQUARES FROM THE MULITPLE REGRESSION MODEL FOR M INCLUDING THE NINE CENSUS REGIONS

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	0.9144551	2.6660525	0.7325
Male III-A Population	9.85805E-07	0.0000043	0.8205
Percentage of Males at the Reserve Center	3.9223890	1.9887836	0.0522
Civilian Wages/Military Wages	-0.3825306	0.1197831	0.0020
Unemployment	-0.1552018	0.1481790	0.2982
Percentage of Prior Service Personnel	-6.2373453	2.9405152	0.0372
Propensity to Commute	0.0425049	0.0167743	0.0133
Census Regions			
East North Central	0.9623399	0.8833887	0.2794
East South Central	-0.7371375	1.1091046	0.5083
Mid-Atlantic	-0.9031297	0.9550033	0.3473
Mountain	-1.8158783	0.9046575	0.0483
New England	-1.2672041	1.1290691	0.2652
Pacific	0.8325207	1.0488923	0.4298
South Atlantic	-0.9086175	1.0473629	0.3884
West North Central	-1.4013373	1.0291937	0.1774
West South Central	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.2538

Model S.S. = 103.3648

Error S.S. = 303.8369

F Value = 1.85

Prob. > F = 0.0466

TABLE 32

PARAMETER SUM OF SQUARES FROM THE MULTIPLE REGRESSION MODEL FOR T INCLUDING THE NINE CENSUS REGIONS

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	-0.5454899	0.3858434	0.1615
Male III-A Population	6.07259E-06	0.0000006	0.3365
Percentage of Males at the Reserve Center	0.7014002	0.2878259	0.0172
Civilian Wages/Military Wages	0.0054210	0.0173356	0.7554
Unemployment	0.0003004	0.0214452	0.9889
Percentage of Prior Service Personnel	-0.4388938	0.4255649	0.3057
Propensity to Commute	0.0072217	0.0024277	0.0039
Census Regions			
East North Central	-0.2368959	0.1278481	0.0678
East South Central	-0.0296627	0.1605147	0.8539
Mid-Atlantic	-0.1412891	0.1382125	0.3099
Mountain	-0.4817537	0.1309262	0.0004
New England	-0.1642686	0.1634041	0.3179
Pacific	-0.0709449	0.1518005	0.6416
South Atlantic	-0.1563604	0.1515792	0.3056
West North Central	-0.2495216	0.1489496	0.0980
West South Central	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.3316

Model S.S. = 3.1570

Error S.S. = 6.3639

F Value = 2.69

Prob. > F = 0.0030

APPENDIX Q

TABLE 33

PARAMETER SUM OF SQUARES FROM THE MULTIPLE REGRESSION MODEL FOR T INCLUDING UNIT TYPE

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	-0.7251097	0.3820098	0.0612
Male III-A Population	9.09502E-07	0.0000006	0.1321
Percentage of Males at the Reserve Center	0.4748144	0.4439196	0.2880
Civilian Wages/Military Wages	0.0065104	0.0173431	0.7084
Unemployment	0.0093589	0.0186847	0.6178
Percentage of Prior Service Personnel	-0.2009020	0.4044274	0.6207
Propensity to Commute	0.0061539	0.0024986	0.0159
Unit Type			
Combat	0.0906429	0.1632313	0.5802
Combat Support	0.1067692	0.1314565	0.4191
Combat Service Support	0.0387408	0.1327390	0.7711
Medical	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.1704

Model S.S. = 1.6224

Error S.S. = 7.8986

F Value = 1.85

Prob. > F = 0.0719

TABLE 34

PARAMETER SUM OF SQUARES FROM THE MULTIPLE REGRESSION MODEL FOR M INCLUDING UNIT TYPE

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > $\lvert t \rvert$
Intercept Term	-2.9060647	2.5470613	0.2573
Male III-A Population	4.93705E-06	0.0000040	0.2191
Percentage of Males at the Reserve Center	4.4396847	2.9598470	0.1375
Civilian Wages/Military Wages	-0.23008624	0.1156354	0.0500
Unemployment	0.0108268	0.1245809	0.9310
Percentage of Prior Service Personnel	-3.9454037	2.6965313	0.1473
Propensity to Commute	0.0304314	0.0166596	0.0714
Unit Type			
Combat	-0.6207960	1.0883493	0.5700
Combat Support	-0.5782258	0.8764898	0.5113
Combat Service Support	0.3493952	0.8850413	0.6940
Medical	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.13769

Model S.S. = 56.0662

Error S.S. = 351.1355

F Value = 1.44

Prob. > F = 0.1863

TABLE 35

PARAMETER SUM OF SQUARES FROM THE MULTIPLE REGRESSION MODEL FOR S INCLUDING UNIT TYPE

Variable	Parameter Estimate	Standard Error of Estimate	Prob. > |t|
Intercept Term	2.9771207	1.2133411	0.0163
Male III-A Population	-3.75635E-06	0.0000019	0.0513
Percentage of Males at the Reserve Center	-0.8058185	1.4099795	0.5692
Civilian Wages/Military Wages	0.0175577	0.0550851	0.7507
Unemployment	-0.0508008	0.0593465	0.3945
Percentage of Prior Service Personnel	-1.1647825	1.2845440	0.3672
Propensity to Commute	-0.0172668	0.0079361	0.0325
Unit Type			
Combat	-0.3742882	0.5184559	0.4724
Combat Support	-0.1889522	0.4175326	0.6521
Combat Service Support	-0.0538954	0.4216063	0.8986
Medical	0.0000000	Base Case	Base Case

Overall Analysis of Variance (ANOVA) statistics are shown below.

R-Square = 0.1459

Model S.S. = 13.6213

Error S.S. = 79.6823

F Value = 1.54

Prob. > F = 0.1486

LIST OF REFERENCES

1. Church, George J. "How Much Is Too Much?" Time Magazine, February 12,1990.

2. Mehay, Stephen L. An Enlistment Supply and Forecasting Model for the U.S. Army Reserve. Technical Report, Fort Sheridan, Illinois: U.S. Army Recruiting Command, July 1989.

3. Wheeler, James O. "The Long, Long Trail to Work." Geographical, Vol. 53, 511-516, May 1981.

4. Monroe, Charles B. and Thomas Maziarz. "American Work-Trip Distances: A Reversal of the Historical Trend." Geography, Vol. 70, 359-362, October 1985.

5. The Journey to Work in the United States: 1979. Special Studies, Series P-23, No. 122, U.S. Department of Commerce, Bureau of the Census.

6. Telephone conversation between Gloria Swieczkowski from the Bureau of the Census and Steven E. Galing on June 20, 1990.

7. Asch, Beth J. Manning The Naval Reserve Force. Center For Naval Analysis, CRM 85-73.10, Alexandria, Virginia, October 1985.

8. Johnson, Laura D. and George W. Thomas. A Maximum Likelihood Method for Characterizing Part-Time Commute Behavior. Naval Postgraduate School, Monterey, California, June 1990.

9. Johnson, Norman L. and Samuel Kotz. Continuous and Univariate Distributions - 2. Palo Alto: Hougton Mifflin Company, 1970.

10. Chambers, John M., William S. Cleveland, Beat Kleiner, and Paul A. Tukey. Graphical Methods for Data Analysis. Wadsworth & Brooks/Cole Publishing Company, 1983, 243-306.

11. Walpole, Ronald E. and Raymond H. Myers. Probability and Statistics for Engineers and Scientists. New York: Macmillan Publishing Co., Inc., 1978.

12. Brooke, A., D. Kendrick and A. Meeraus. GAMS: A User's Guide. The Scientific Press, 1988.

13. Bain, Lee J. and Max Engelhart. Introduction to Probability and Mathematical Statistics. Boston: Duxbury Press, 1987.

14. Box, G.E.P. and D.R. Cox. An Analysis of Transformations. JRSS, 1964, 2, 211-250.

15. Army Regulation 601-210. Regular Army and Army Reserve Enlistment Program. Headquarters Department of the Army, Washington, DC, 1 December 1988.

16. County and City Data Book 1988. A Statistical Abstract Supplement. U.S. Department of Commerce. Bureau of the Census, Washington, DC.

17. State and Metropolitan Area Data Book 1986. A Statistical Abstract Supplement. U.S. Department of Commerce. Bureau of the Census, Washington, DC.

18. Gibra, Issac. Probability and Statistical Inference for Scientists and Engineers. Prentice-Hall, Inc., 1973.

19. SAS User's Guide: Statistics. Version 5 Edition. SAS Institute Inc., Cary, NC, 1985.

BIBLIOGRAPHY

Army Regulation 601-210. Regular Army and Army Reserve Enlistment Program. Headquarters Department of the Army, Washington DC, 1 December 1988.

Asch, Beth J. Manning The Naval Reserve Force. Center For Naval Analysis, CRM 85-73.10, Alexandria, Virginia, October 1985.

Bain, Lee J. and Max Engelhardt. Introduction to Probability and Statistics. Duxbury Press, 1987.

Bazaraa, Mokhtar S. and C. M. Shetty. Nonlinear Programming Theory and Algorithms. John Wiley & Sons, 1979.

Box, G.E.P. and D.R. Cox. An Analysis of Transformations. JRSS, 1964, No. 2, 211-250.

Brooke, A., D. Kendrick and A. Meeraus. GAMS: A User's Guide. The Scientific Press, 1988.

Chambers, John M., William S. Cleveland, Beat Kleiner, and Paul A. Tukey. Graphical Methods for Data Analysis. Wadsworth & Brooks/Cole Publishing Company, 1983, 243-306.

Church, George E. "How Much Is Too Much?" Time Magazine, February 12, 1990.

County and City Data Book 1988. A Statistical Abstract Supplement. U.S. Department of Commerce. Bureau of the Census, Washington, DC.

Gibra, Issac. Probability and Statistical Inference for Scientists and Engineers. Prentice-Hall, Inc., 1973.

Goldberg, Matthew S. and Stanley A. Horowitz. Constraints On Expansion of the Reserve Forces. Institute for Defense Analysis, IDA Paper P-2358, April 1990.

Gordon, Peter, Ajay Kumar and Harry W. Richardson. "The Influence of Metropolitan Spatial Structure on Commuting Time." Journal of Urban Economics, 1989, Vol. 26, 138-151.

Hamilton, Bruce W. "Wasteful Commuting." Journal of Political Economy, 1982, Vol. 90, No. 5, 1035-1051.

Izraeli, Oded and Thomas R. McCarthy. "Travel Time and Modal Choice Among SMSA's." Journal of Transportation Economics and Policy, 1985, Vol. 19, No. 2, 139-160.

Johnson, Laura D. and George W. Thomas. A Maximum Likelihood Method for Characterizing Part-Time Commute Behavior. Naval Postgraduate School, Monterey, California, June 1990.

Johnson, Norman L. and Samuel Kotz. Continuous and Univariate Distributions - 2. Palo Alto: Hougton Mifflin Company, 1970.

Madden, Janice Fanning. "Urban Wage Gradients: Empirical Evidence." Journal of Urban Economics, Vol. 18, 291-301, 1985.

Mehay, Stephen L. An Enlistment Supply and Forecasting Model for the U.S. Army Reserve. Technical Report, Fort Sheridan, Illinois: U.S. Army Recruiting Command, July 1989.

Monroe, Charles B. and Thomas Maziarz. "American Work-Trip Distances: A Reversal of the Historical Trend." Geography, Vol. 70, 356-362, October 1985.

Morrison, Peter A. and Allan Abrahamse. Is Population Decentralization Lengthening Commuting Distances?" Rand Corporation, N-1934-NICHD, Santa Monica, California, December 1982.

Neter, John, William Wasserman, and Michael H. Kutner. Applied Linear Regression Models. Second Edition. Richard D. Irwin, Inc., 1989.

Road Atlas. Rand McNally and Company, 1986.

SAS User's Guide: Statistics. Version 5 Edition. SAS Institute Inc., Cary, NC, 1985.

State and Metropolitan Area Data Book 1986. A Statistical Abstract Supplement. U.S. Department of Commerce. Bureau of the Census, Washington, DC.

The Journey to Work in the United States: 1979. Special Studies, Series P-23, No. 122. U.S. Department of Commerce. Bureau of the Census, Washington, DC.

Walpole, Ronald E. and Raymond H. Meyers. Probability and Statistics for Engineers and Scientists. New York: Macmillan Publishing Co., Inc., 1978.

Ward, William F. "Top-Notch Training and Equipment Raise Reserve Readiness Levels." Army, 100-107, October 1989.

Westcott, Diane N. "Employment and Commuting Patterns: a Residential Analysis." Monthly Labor Review, 3-9, July 1979.

Wheeler, James O. "The Long, Long Trail to Work." Geographical, Vol. 53, 556-561, May 1981.

Telephone conversation between Gloria Swieczkowski from the Bureau of the Census and Steven E. Galing on June 20, 1990.

"The Journey to Work in the United States." Special Studies, Series P-23, No. 122, December 1982, p. 4.